# VISUAL SEARCH AND IMAGE QUALITY

Harry L. Snyder, et al

Virginia Polytechnic Institute and State University

Prepared for:

Aerospace Medical Research Laboratory

October 1974

# NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Organizations and individuals receiving announcements or reports via the Aerospace Medical Research Laboratory automatic mailing lists should submit the addressograph plate stamp on the report envelope or refer to the code number when corresponding about change of address or cancellation.

Do not return this copy. Retain or destroy.

Please do not request copies of this report from Aerospace Medical Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22151

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DODD 5230.0. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

JULIEN M. CHRISTENSEN, Ph.D.
Director, Human Engineering Division
6570 Aerospace Medical Research Laboratory

AIR FORCE/56780/18 February 1975 — 150

AD-A008 007

Security Classification

**DOCUMENT CONTROL DATA - R & D**

*(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)*

1 ORIGINATING ACTIVITY *(Corporate author)*

Department of Industrial Engineering & Operations Research, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3 REPORT TITLE

VISUAL SEARCH AND IMAGE QUALITY

4 DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

INTERIM REPORT 6/15/71 - 3/15/73

5 AUTHOR(S) *(First name, middle initial, last name)*

Harry L. Snyder, Ph.D. William S. Beamon
Robin L. Keesee James R. Aschenbach, M.S.

| 6 REPORT DATE | 7a. TOTAL NO. OF PAGES | 7b. NO OF REFS |
|---|---|---|
| October 1974 | 120 | 33 |

| 8a. CONTRACT OR GRANT NO | 9a. ORIGINATOR'S REPORT NUMBER(S) |
|---|---|
| F33615-71-C-1739 | |
| b. PROJECT NO 7183 | |
| c. Task No. 718300 | 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* |
| d. Work Unit No. 71830004 | AMRL-TR-73-114 |

10 DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

| 11 SUPPLEMENTARY NOTES | 12. SPONSORING MILITARY ACTIVITY |
|---|---|
| | Aerospace Medical Research Laboratory<br>Aerospace Medical Division, AFSC<br>Wright-Patterson AFB, OH 45433 |

13. ABSTRACT

Several experiments were conducted to evaluate alternate unitary measures of video line-scan system image quality. A metric based upon the Modulation Transfer Function of the imaging system was derived, with emphasis placed upon the photometric properties of the system. This metric was shown to predict well the average effects of several imaging system parameters upon the ability of observers to extract information from both dynamic and static images. In attempting to predict the ability of observers to acquire specific targets in an air-to-ground search task, however, other target and background parameters become very important, and such image quality measures must therefore be refined. Relationships among alternate measures of line-scan image quality were discussed, and a conceptual model was presented for combining system noise, raster interference, scene content, and the visual requirements of the observer.



DD FORM 1 NOV 65 1473

Security Classification

| 14. KEY WORDS | LINK A | | LINK B | | LINK C | |
|---|---|---|---|---|---|---|
| | ROLE | WT | ROLE | WT | ROLE | WT |
| Human engineering<br>Image quality<br>Target acquisition<br>Television | | | | | | |

## PREFACE

This study was initiated by the Performance Requirements Branch, Human Engineering Division, of the Aerospace Medical Research Laboratory. The research was conducted by the Department of Industrial Engineering and Operations Research, of Virginia Polytechnic Institute and State University of Blacksburg, Virginia 24061, under Air Force contract F33615-71-C-1739. Dr. Harry L. Snyder was the principal investigator for Virginia Polytechnic Institute and State University. Mr. James L. Porterfield was the Technical Monitor for the Aerospace Medical Research Laboratory. The report covers research performed between June 1971 and March 1973.

TABLE OF CONTENTS

# SECTION I

# INTRODUCTION

Remote sensing, line-scan imaging systems have been widely used by the U.S. Air Force in Southeast Asia, and, no doubt, will be used in the future, if necessary, in other locales. One major problem in designing such systems, for example low-light-level television or infrared, has been the lack of criteria by which one can predict, or against which one can evaluate, the performance of the operator viewing the display. While numerous studies have investigated parts of this problem, a comprehensive basis for specifying line-scan display image quality in relation to human operator performance has not yet evolved.

This report summarizes the first phase of a research program designed to determine the relationships between target recognition performance measures and the more promising indices of image quality. Emphasis is placed upon predicting the performance of a given line-scan system, including its operator, in recognizing both diverse and specific targets. The objective of the first phase of this program was to provide experimental data comparing different alternate measures of image quality, and to identify the image quality metric showing the most promise for predicting operator target recognition performance with a line-scan display.

The second phase, already begun, is intended to determine the limits of generalization of the recommended unitary measure of image quality. In addition, the second phase will provide eye-movement data to evaluate visual search patterns and parameters as a function of image quality, in an attempt to obtain a better understanding of how image quality affects the target recognition process.

The third phase of the research program will develop a set of system design criteria for predicting operator performance as a function of image quality. It shall also include a model of visual search as related to mission and system parameters.

## NEED FOR A UNITARY MEASURE OF IMAGE QUALITY

The problem of specifying the image quality of line-scanning systems received increased attention in the early 1960's with the advent of low-light-level television and infrared imaging systems for reconnaissance and strike aircraft. In addition, the possibility of both manned and unmanned exploration of the lunar surface spurred interest in improving the telemetering of image data. The need to better understand image quality became particularly apparent when it was realized that digitizing of the video signal for transmission introduced a "new" form of image noise ("striping"). As a result of these several more-or-less simultaneous needs and interests, research into the nature of line-scan image quality and its effect upon image interpretability was begun about 1961, and has continued through the present.

During the past 12 years, over 300 laboratory and analytical studies have been performed to assess the relationship between variation in line-scan display image parameters and observer performance. Unfortunately, critical reviews of these studies indicate that cross-study comparisions are virtually impossible. For example, variations in specific system design parameters or in the techniques of synthetically manipulating image quality are often incompletely controlled, resulting in indeterminant concomitant variation in other potentially relevant factors. Table 1 lists some of the experimental variables which have been shown to significantly affect the operator's information extraction (e.g., target acquisition) performance. Note that individual experiments tend to examine the effects of one or two, rarely three, such variables. Due to the inherent interaction among many of these variables, quantitative combination of the results is hazardous even in the presence of good experimental control. In the absence of such control, any a posteriori combining of the results is probably impossible.

Recently, various investigators have directed their efforts toward developing, either mathematically or experimentally, a summary measure of image quality which both takes into account the numerous parameters of a line-scanning system and predicts its performance, usually in terms of some objective measure of operator performance. Because such investigators have come from diverse backgrounds and

TABLE I

SOME OF THE VARIABLES AFFECTING OBSERVER TARGET RECOGNITION PERFORMANCE

| Atmospheric | Scene |
|---|---|
| Aerosol Content | Target Characteristics |
| Cloud Cover | Background Characteristics |
| Illumination Level | Terrain Masking |
| **Sensor** | Clutter Level |
| Bandwidth | **Display** |
| Number of Scan Lines | Luminance |
| Field of View | Size |
| Field/Frame Rate | Number of Scan Lines |
| Aspect Ratio | Contrast |
| S/N Level | Scene Movement |
| Integration Time | Dynamic Range |
| **Image Processing** | Gamma |
| Edge Enchancement | S/N Level |
| Gamma | Aspect Ratio |
| Spatial Filtering | |

have varying interests, these several measures of image quality are couched in different terms, and are derived in decidedly different ways. Although at a first glance some of these measures appear quite different, as will be shown later, they may be quite similar in terms of final prediction. To relate the research of this program to these various measures of image quality, the several alternate candidates are summarized in the following paragraphs.

## ALTERNATE MEASURES OF IMAGE QUALITY

Previous research pertinent to the specification of line-scan image quality has come from two totally separated areas of commercial activity - the television industry and silver halide photography. Each of these will be discussed briefly.

### Television-Related Research

A television system, not unlike most present-day reconnaissance line-scan systems, has a finite aperture response which causes it to transmit the contrast of grids or bars less well as the grid elements or bars move closer together. A typical television response curve has the form shown in figure 1.

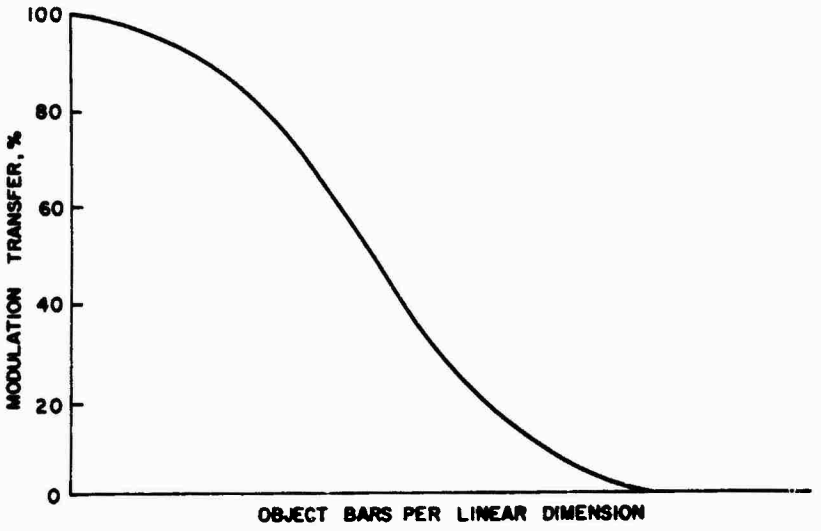



Figure 1. Typical TV System Sine-Wave Response

Using a sine-wave target modulation, rather than the standard square-wave target (as represented for example by the USAF tri-bar target), one can evaluate the modulation transfer of each element (e.g., lens, preamplifier, display) in a video system, and then multiply these modulation transfer curves to obtain the overall system response. An example is shown in figure 2. Employing this technique, the system response for a given target size is simply the point-by-point product of the component response curves.

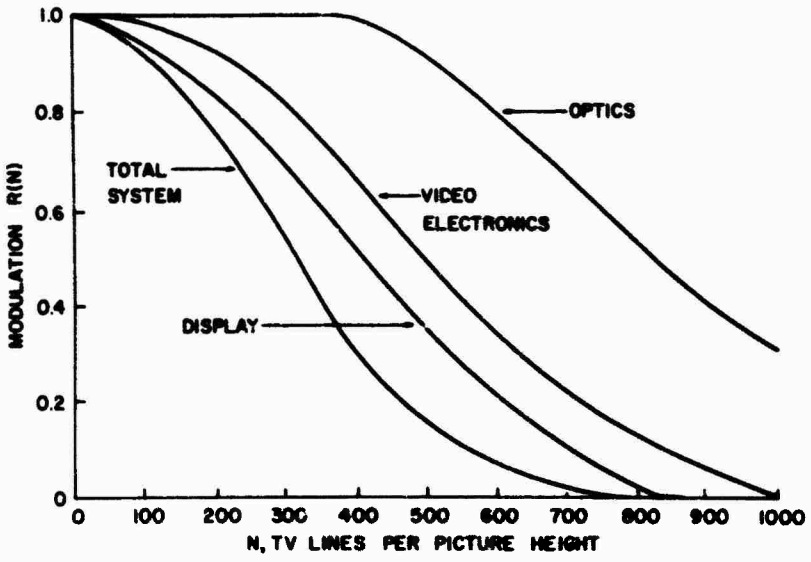



Figure 2. Cascading of Components to Obtain System MTF. The parameter TV Lines per Picture Height is the number of half sinusoidal cycles imaged upon the sensor across its smaller dimension, conventionally the vertical dimension in the 4:3 (horizontal: vertical) aspect ratio, where the raster lines are horizontally oriented.

Given knowledge of this response curve, it is often convenient (ref. 1) to consider the quality of the television image as proportional to the Equivalent Passband, $N_e$, the passband of an equivalent rectangular noise spectrum with an abrupt cutoff (at spatial frequency $N_e$) which passes the same <u>total</u> sine-wave energy as the actual spectrum. This concept is illustrated in figure 3. It should be noted that the sine-wave response is one dimensional, but that $N_e$ is the two-dimensional aperture response of the system, and therefore is determined from the square of the one dimensional sine-wave response:

$$N_e = \int_0^\infty [R(N)]^2 \, dN \qquad (1)$$

where R(N) is the percent response, and
N is the spatial frequency in TV lines/picture height.

This summary measure has been derived and pioneered by Schade, and has been accepted by many in the industry for years. For usage in performance prediction of present day reconnaissance systems, however, it appears to have one liability; namely, that it does not take into account the varying noise levels which a system might have as, for example, the detector irradiance level changes with changes in scene illumination.

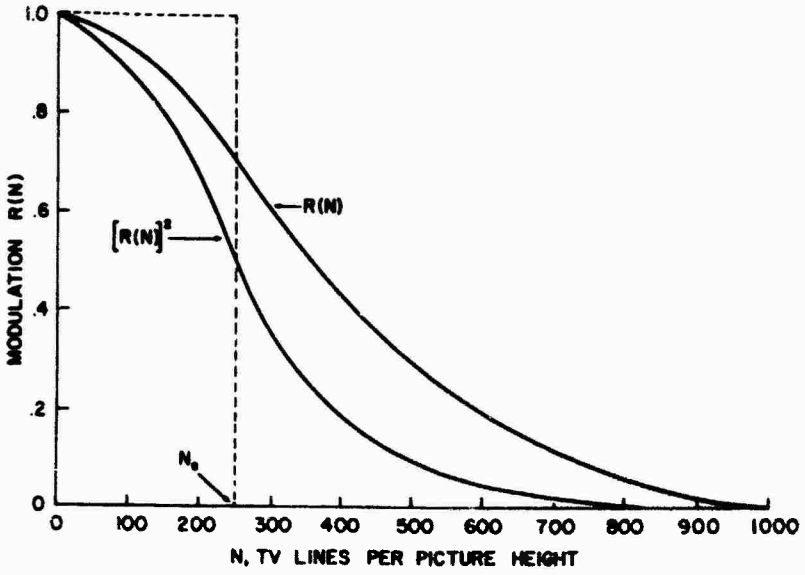



Figure 3. Noise Equivalent Passband, $N_e$.

Using the analyses of Schade (e.g., ref. 1) as background, Rosell (ref. 2) has developed an approach for analyzing television systems which gets closer to the human observer's visual capability. Rosell's approach is to relate all system parameters to the analytically derived signal-to-noise ratio at the display ($SNR_D$). Then, assuming the human observer requires an $SNR_D$ of approximately 1.2 to have a 50% chance of detecting a target, system tradeoffs are made to achieve this or some other value of $SNR_D$. Many laboratory studies have been performed to establish the probability of detection of gratings and solid rectangles as a function of $SNR_D$. Observer confidence levels, task loading, ambient environments, dynamic scenes, target textural characteristics, and other factors have not been considered. While this concept shows promise, empirical human performance data are required to make it more generally acceptable and useful.

There are many variants of the $SNR_D$ concept, depending upon whether one assumes the limitations in the line-scan system to be, for example, photon limited, preamplifier limited, display limited, etc. For purposes of discussion, however, an elementary calculational formula is given by Rosell (ref. 3, p. 18):

$$SNR_D = \left[at \cdot \Delta f_V / A\right]^{1/2} \frac{Ci_{max}}{(2-C)\, e\Delta f_V\, i_{max}^{1/2}} \tag{2}$$

$$= \left[(a/A) \cdot t \cdot \Delta f_V\right]^{1/2} SNR_V \tag{3}$$

where $SNR_D$ = signal-to-noise ratio at the display
$a$ = area subtended by target at photosurface
$A$ = total area of photosurface
$t$ = integration time of eye, assumed to be between 0.1 and 0.2 sec.
$\Delta f_V$ = video bandwidth, in hertz
$C$ = target contrast
$i_{max}$ = maximum photocurrent
$e$ = charge of an electron
$SNR_V$ = signal-to-noise ratio in the video

As Rosell points out, the same value of $SNR_D$ is obtained for different size targets if the $SNR_V$ at threshold varies inversely with the solid angular subtense of the target, $a$. In his experiments, the value of $SNR_D$ is essentially constant within the limits of the spatial integration capability of the eye, assuming optimum viewing distance for a given display size. The desirability of this model lies in its derivation directly from the parameters of the sensor tube and camera being evaluated, and in its utility in providing tradeoff data for all parameters of any type of sensor system.

As an example of the development of this model and its application to specific system parameters, one might consider one variation of the $SNR_D$ formula (ref. 3):

$$SNR_D = \frac{(\alpha\, t)^{1/2}}{N} \cdot \frac{C\, R_{SQ}(N)\, G\, i_s/e_v e_h}{\left[G^2\, e\, i_s/e_v e_h + I^2_{PA}/\Delta f\right]^{1/2}} \qquad (4)$$

in which $SNR_D$ = display signal-to-noise ratio

$\alpha$ = picture aspect ratio

$t$ = integration time of the eye, assumed to be $\simeq$ 0.2 sec.

$N$ = number of resolution elements (e.g., TV lines) per picture height

$C$ = image contrast

$R_{SQ}(N)$ = system response factor at N

$G$ = signal amplification

$i_S$ = photocurrent

$e_v$ = vertical scan efficiency

$e_h$ = horizontal scan efficiency

$e$ = charge of an electron

$I^2_{PA}$ = mean square preamplifier noise

$\Delta f$ = video bandwidth

This equation serves as the basis for evaluating many conceivable line-scan imaging systems. By determining apparent target contrast (C) as a function of slant range, atmospheric effects, and inherent target contrast according to well-established relationships (ref. 4), it is possible to calculate the slant range at which $SNR_D = 1.2$, the assumed value for which the probability of target detection equals 0.50.

Several generalizations can be made from examination of this equation. From the first term on the right side of the equation, as the number of resolution elements, N, decreases (or, equivalently, as target spatial frequency decreases), the $SNR_D$ will increase. Typically, $\alpha$ (picture aspect ratio) and t (integration time of the eye) remain constant. Increases in C (image contrast or target apparent contrast), $R_{SQ}(N)$ (system square-wave modulation response), and G (system gain) all serve to increase $SNR_D$. These relationships are similar to

those which will be subsequently expressed for the MTFA concept if G is considered similar to photographic gamma. In the denominator of the right-hand terms, increases in G also increase noise, causing a reduction of $SNR_D$. As will be discussed in Section VII of this report, the $SNR_D$ concept is very similar to other measures, and, under some circumstances, is mathematically equivalent.

Photographic Research

Although there have been several studies investigating relationships between subjective image quality and such physical measures of the photographic image as limiting resolution, granularity, and acutance (e.g., references 5-7), it was not until 1965 that a promising unitary measure of photographic image quality was suggested. That measure is typically referred to today as the Modulation Transfer Function Area (MTFA).

Originally developed analytically by Charman and Olin (ref. 8), who termed it the threshold quality factor, and later investigated empirically and renamed by Borough, et al. (ref. 9), the MTFA concept has been employed in two photographic experiments, which have demonstrated that it relates strongly to the ability of image interpreters to obtain critical information from reconnaissance photographic imagery. In its original form, the MTFA was proposed as a unitary measure of photographic image quality which contains "the cumulative effect of the various stages of the atmosphere-camera-emulsion-development-observation process, the 'noise' introduced in the perceived image by photographic grain, and the limitations imposed by the physiological and psychological systems of the observer" (ref. 8, p. 385).

The MTFA is derived in such a manner as to make use of the Modulation Transfer Function (MTF) of the imaging system, thereby retaining the analytical convenience of component analysis based upon the sine-wave response characteristic, the same response characteristic which forms the basis of the $N_e$ and $SNR_D$ measures. In addition, the MTFA attempts to take into account other variables critical to the imaging and interpreting problem, such as exposure, the characteristic curve, granularity, the human observer visual capabilities and limitations,

and the nature of the interpretation task. (For the electro-optical system, the first three of these variables can be considered analogous to detector irradiance level, gamma, and noise, respectively.)

Figure 4 shows that the MTFA is the area bounded by the imaging system MTF curve and the detection threshold curve of the total system, including the eye. The MTF curve for the imaging system is obtained in the conventional manner, while the detection threshold curve requires several assumptions regarding the human operator. Specifically, it is assumed that the viewing conditions are optimum, and that threshold detection of any target in the image is a function of the target (image) contrast modulation, the noise in the observer's visual system, and the noise in the imaging system exclusive of the observer. It should be noted that the crossover of the two curves in fig. 4 determines the conventional limiting resolution of the system for a sine-wave target.

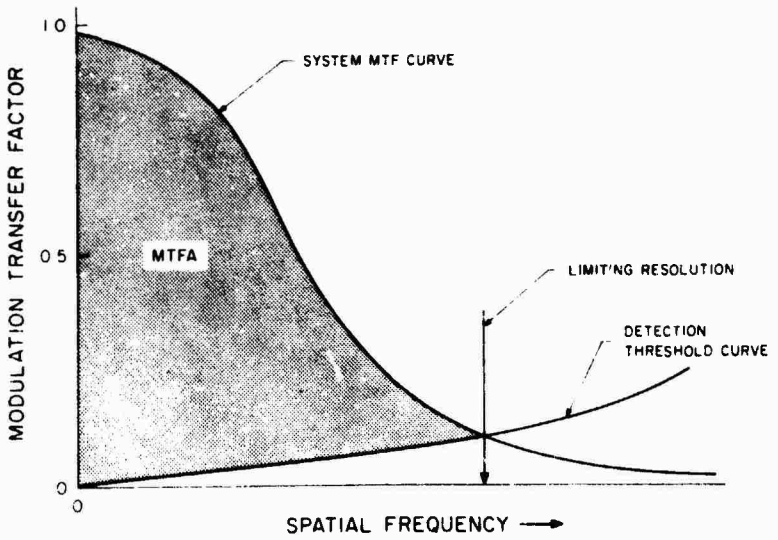



Figure 4. Modulation Transfer Function Area (MTFA).

At low spatial frequencies, the threshold detection curve is dependent upon the limiting properties of the human visual system, as shown in fig. 5. At higher spatial frequencies, the effect of imaging system noise becomes important. For the photographic image, this imaging system noise is equivalent to granularity, which is assumed to be Gaussian. It is assumed further that the eye's modulation threshold is 0.04, so that a target image modulation of 0.04 must be realized

for the target to be detected, regardless of the modulation of the target object.

Figure 5 illustrates the normalized detection threshold curve (ref. 9), which must be adjusted both vertically and horizontally for a specific set of conditions. First, the curve is positioned vertically by increasing the normalized ordinate scale by $\frac{M_t(N)}{M_o}$, where $M_t(N)$ is the normalized value as shown in fig. 5, and $M_o$ is the target contrast modulation. The lower portion of the threshold curve (at the lower spatial frequencies) is also adjusted by the system gamma, which, if greater than unity, enhances the modulation recorded at the display (e.g., film) so that the minimum detectable threshold modulation decreases by $\frac{0.04}{\gamma}$

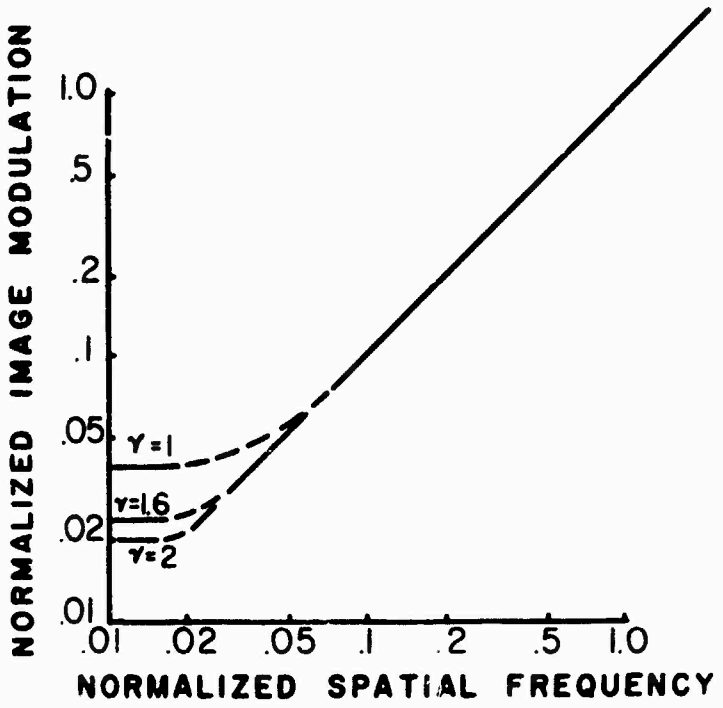



Figure 5. The MTFA Generalized Detectability Threshold.

Next, the detection threshold curve is positioned horizontally by multiplying the scale of the abscissa in fig. 5 by $\frac{2}{C\sigma(D)}$, where C is an empirically-derived constant (0.03 for fine-grained films and 0.04 for coarser grained films, ref. 10), and $\sigma(D)$ is equal to the rms granularity measured with a 24-micron scanning aperture.

Algebraically, the detection threshold curve for a photographic system is therefore (ref. 9):

$$M_t(N) = 0.034 \left[\frac{dD}{d(\log_{10}E)}\right]^{-1} \left[0.033 + \sigma(D)^2 N^2 S^2\right]^{1/2} \qquad (5)$$

in which N = any spatial frequency, in lines per millimeter,

0.034 = a theoretically derived constant, (1)

D = mean film density,

E = exposure,

0.033 = a theoretically derived constant, (1)

$\sigma(D)$ = rms granularity for a 24μ scanning aperture,

S = signal-to-noise ratio necessary for threshold viewing, assumed to be about 4.5 (ref. 11), and

$\frac{dD}{d(\log_{10}E)}$ = film characteristic slope, including the effects of development.

When the MTF curve and the detection threshold curve are plotted on log-log coordinates (ref. 9), the expression for the MTFA becomes:

$$\text{MTFA (log-log)} = \int_{\log N_o}^{\log N_1} (\log T_N)\, d \log N - \int_{\log N_o}^{\log N_1} \log\left(\frac{M_t(N)}{M_o}\right) d \log N$$

$$= \int_{\log N_o}^{\log N_1} \left(\log \frac{M_o T_N}{M_t(N)}\right) d \log N \qquad (6)$$

where $N_o$ = the low spatial frequency limit, in lines/millimeter,

$N_1$ = the spatial frequency at which the MTF curve crosses the detection threshold curve (limiting resolution),

$T_N$ = the MTF value at spatial frequency N,

$M_o$ = the target contrast modulation

$= \frac{\text{luminance of target - luminance of background}}{\text{luminance of target + luminance of background}}$, and

$M_t(N)$ = the normalized detection threshold curve value, as taken from fig. 5.

---

(1) For derivation, see Charman and Olin (ref. 8). Generation of these values is considered unimportant in the present context.

When the MTF curve and the detection threshold curve are plotted on linear coordinates, the area of interest is given by (ref. 9):

$$\text{MTFA (linear)} = \int_0^{\nu_1} \left( T_N - \frac{M_t(N)}{M_o} \right) dN. \qquad (7)$$

The linear form computation utilizes no lower frequency cutoff, whereas the log-log formulation employs an arbitrary cutoff at, say, 10 lines/millimeter. The reason for this difference is simply because the log-log plot integration would place an inappropriately large weight upon integration over the lower spatial frequencies were this cutoff eliminated. The nature of the linear plot avoids the need for such an arbitrary cutoff.

It might also be noted, parenthetically, that the detection threshold curve, as described here, is akin to such concepts as contrast sensitivity (ref. 12), sine-wave response (refs. 13, 14, 15), and demand modulation function (ref. 16).

Because of the problems associated with deriving an equivalent expression for the MTFA of a raster-scan display, no discussion of that subject is contained here. Rather, such discussion is in Sections III and VII.

To date, two empirical evaluations of the MTFA concept have been conducted, both using photographic emagery. In the first study (ref. 9), an attempt was made to relate MTFA to subjective estimates of image quality obtained from a large number of trained image interpreters. In the second of these experiments, information-extraction performance data and subjective estimates of image quality were obtained, and both measures were compared with the MTFA values of the imagery. While it is desirable from an operational viewpoint to have a quick judgment of subjective image quality to serve as an indicator of the quality of any source of imagery for, say, rapid screening purposes, the critical measure of quality of any imaging system is the ability of the observer to perform the required information extraction tasks.

In the first study to evaluate MTFA (ref. 9), the purpose was to determine whether a strong relationship existed between MTFA and subjective image quality. Nine photographic reconnaissance negatives were used as the basis for laboratory-controlled manipulation of image quality. Each of the scenes was printed in 32 different MTFA variants, determined by 4 different MTF's, 3 levels of granularity, and 3 levels of contrast, as illustrated in figure 6. Four cells of the matrix were deleted because their MTFA values corresponded to others in the 32-celi matrix. The MTF curves are illustrated in figure 7.

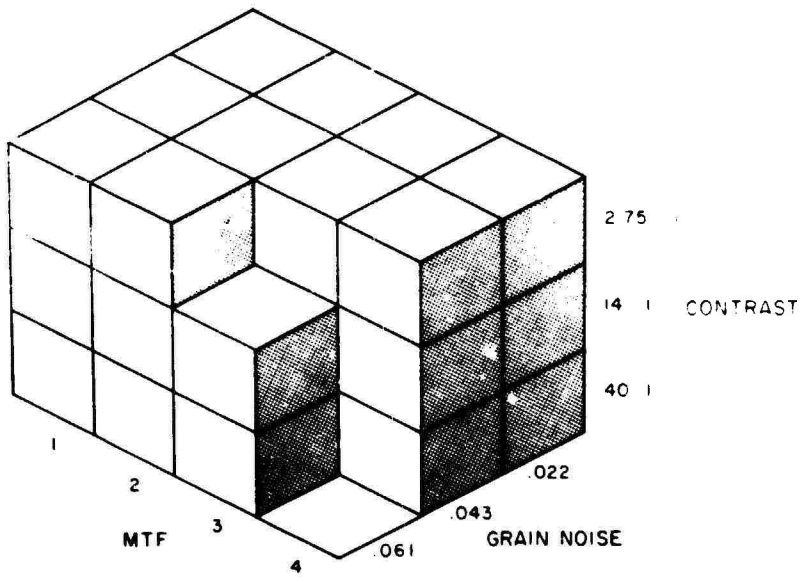



Figure 6. Production of 32 MTFA Values.

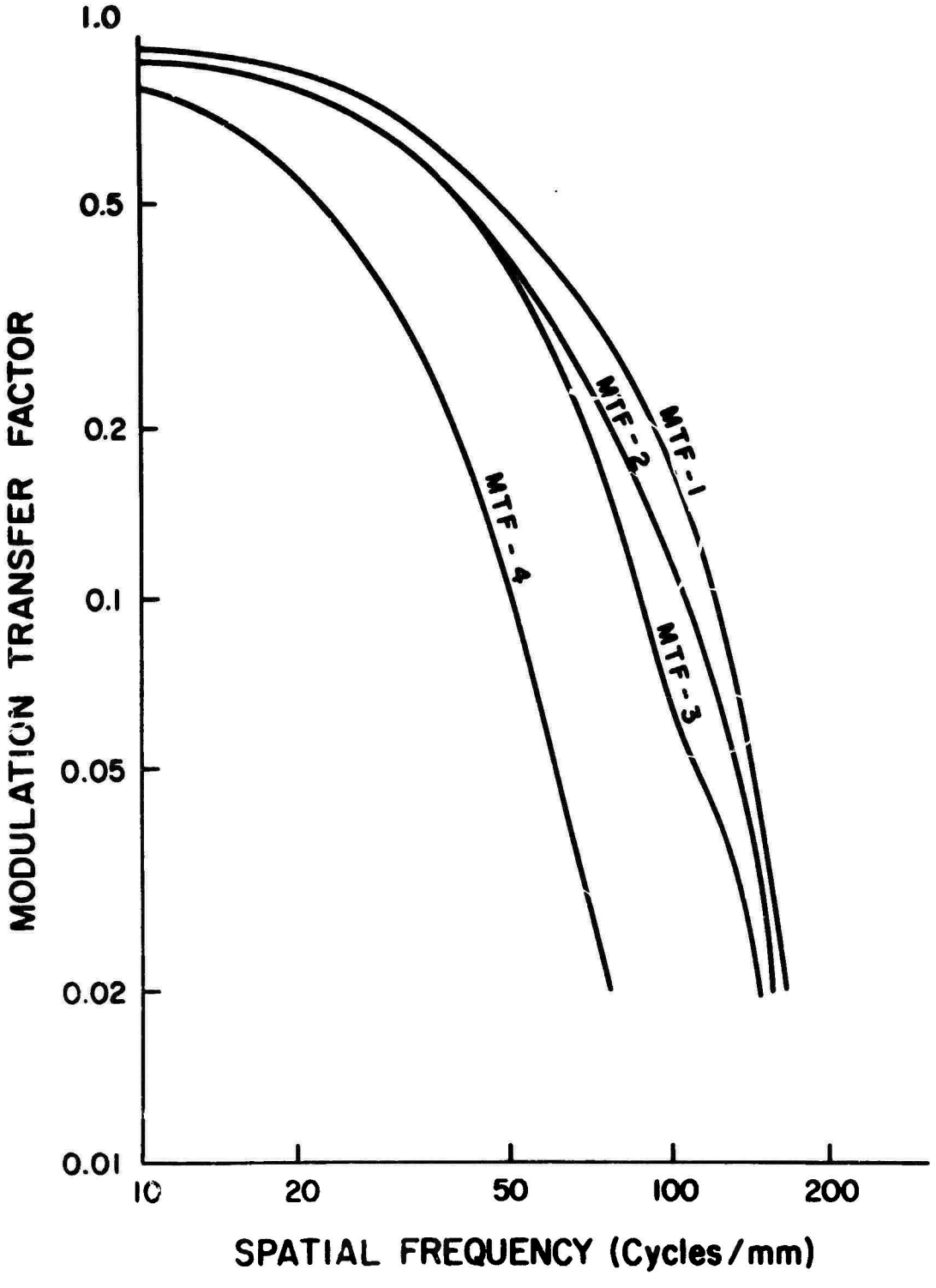



Figure 7. Average Modulation Transfer Curves, from ref. 9.

The resulting 288 transparencies (9 scenes by 32 variants/scene) were used in a partial paired-comparison evaluation. The subjects, 36 experienced photointerpreters, were asked to select the photo of each pair which had the best quality for extraction of intelligence information.

Correlations were obtained between the subjective image quality rating (derived from the paired-comparisons data) for each of the 32 variants and several physical measures of image quality. Most important to this discussion is the obtained mean product-moment correlation of 0.92 between MTFA (linear) and subjective image quality, which indicates that MTFA is strongly related to subjective estimates of image quality.

The next experiment, by Klingberg, Elworth, and Filleau (ref. 17), examined the relationship between objectively measured information-extraction performance and the MTFA values. As a check on the results of Borough, et al., Klingberg, et al. also obtained subjective estimages of image quality, so that all three inter-correlations were evaluated.

The imagery used for this experiment was the same as that used by Borough, et al. (ref. 9). A group of 384 trained military photointerpreters served as subjects. Each subject was given one variant of each of the nine scenes, and was asked to (1) rank the image on a 9-point interpretability scale using utility of image quality for information extraction as the criterion, and (2) answer each of 8 multiple choice questions dealing with the content of the scene. The interpretability scale values were used to develop a subjective image quality measure for the 288 images, while scores on the multiple-choice interpretation questions were used to measure information extraction performance.

Figure 8 shows the scattergram between information extraction performance and MTFA for the 32 MTFA values. The resulting correlation, averaged across the nine scenes, is -0.93. (The minus value is due to the use of number of errors as a measure, which is inversely related to MTFA).

In addition, the correlation between performance and subjective ranking was 0.96, while the correlation between MTFA and subjective rank, replicating the result of Borough, et al. was 0.97. Thus, the MTFA concept, as applied to photographic imagery, is an excellent predictor of both subjective image quality and the measured performance of trained image interpreters to perform an operational task.

## OVERVIEW OF THE PHASE ONE RESEARCH

During the first phase of this research program several experiments were conducted (1) to develop appropriate experimental and measurement techniques which define the physical measures of image quality pertinent to line-scan displays, (2) to determine the relationship between such physical measures and operator performance, both from static and dynamic imagery, and (3) to evaluate the extent to which such metrics predict target acquisition performance for specific targets, rather than predict average performance across a large sample of targets. The following sections of the report summarize these results. Detailed emphasis is placed upon the equipment and techniques employed in the measurements in the hope that some commonality of results may be shared with other investigators.

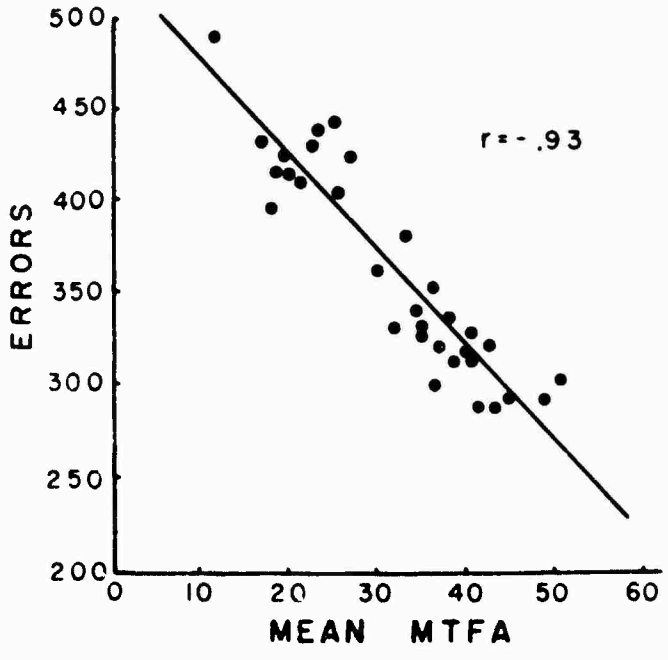



Figure 8. Scattergram of Information-Extraction Performance vs. MTFA.

## SECTION II

## EQUIPMENT AND MEASUREMENT TECHNIQUES

The generalized system employed for this research is illustrated in figure 9 in block-diagram form. A variable-parameter television system is used to display the image of either static or dynamic imagery. An observer, seated before the monitor, responds when he is able to recognize the object of interest on the display in accordance with some set of instructions. His response, indicated by some combination of switch closure and verbal description, is recorded on either appropriate chart recorders or by a printout on a counter/printer.

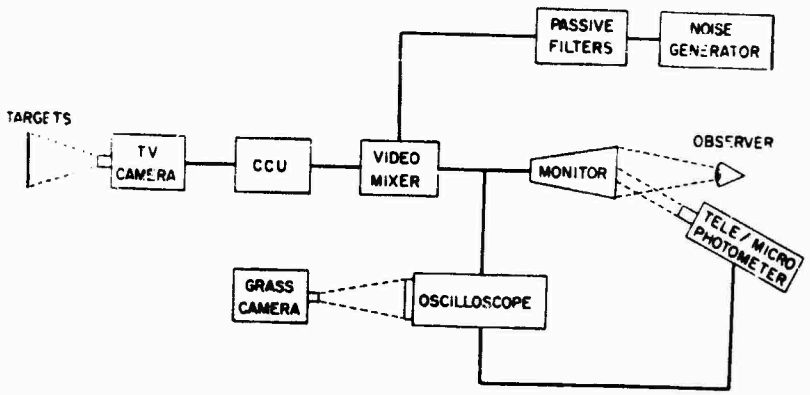



Figure 9. Multiparameter Video System Block Diagram.

Quantitative measurement of the video signal is made electrically in the video chain, and photometrically at both the imagery input and the monitor output. Thus, the image to which the observer responds is quantified in electrical and photometric terms.

### VARIABLE-PARAMETER TELEVISION SYSTEM

The television camera, a Cohu Model 6100, contains an 8607A 1-inch vidicon and has a specified video bandwidth of 32 MHz (-3 dB point). It can be driven by the camera control unit, a Cohu Model 6900, at any line rate from

525 lines per frame to 1225 lines per frame. The camera control unit (CCU) has a comparable specified bandwidth, although measurements made on the system indicate an electrical bandwidth well in excess of 35 MHz. Limiting resolution, as specified by the manufacturer, is approximately 1100 lines per picture height, center resolution, when the full 32-MHz bandwidth is used in conjunction with the 1225-line scan rate. All scan rates are in a 30 frame per second, 2:1 positive interlace format.

By replacing the vidicon with a Cohu vidicon simulator, sweeping the input with a 0-50 MHz sweep generator, and comparing the input sweep amplitude with the output sweep amplitude, taken at various points in the video channel, one can measure the electrical response of the video channel. Results of these measurements show that the sine-wave frequency response of the system is well in excess of 35 MHz, with a fairly smooth rolloff at the higher frequencies.

Several resistors, capacitors, and coils are replaced or tuned in order to achieve a given bandwidth rolloff. These changes, in both the camera video boards and the CCU video board, are made with the sweep generator to assure that no spurious response is inserted into the system.

Similarly, line-rate changes require the changing of a single strip on the CCU's sync generator card, plus adjustment of a variable choke on the camera. These changes present no problem, if made carefully with oscilloscope monitoring of the sync and blanking waveforms.

VARIABLE-PARAMETER MONITOR

The monitor employed in all the experiments described in this report is a Conrac RQA-17, with a P4 phosphor. This monitor is designed to accept all line rates, field rates, and frame rates. External controls include the conventional brightness, contrast, focus, horizontal size, vertical size, etc., plus rear panel adjustments for sync (internal, external), DC restoration, vertical linearity, and horizontal linearity. As shall be noted later, the photometric response of this monitor was less than that desired, although

the measured electrical video bandwidth, from its preamplifier, was in excess of 35 MHz.

## IMAGE QUALITY MANIPULATION

Experimental manipulation of image quality was obtained to match the parametric deterioration which might be expected under operational conditions. Specifically, it was necessary to vary line rate, video bandwidth, noise amplitude, noise passband, and any characteristics of the image peculiar to the operational situation, such as target type, target/background contrast, image scale, rate of image motion through the field of view, etc. To achieve this objective, the variables of interest were categorized into (1) television system variables and (2) mission input variables.

Television system variables were manipulated in the TV system, with the line rate and bandwidth of the video channel being varied as described above. Noise was inserted by taking the output of a General Radio Model 1383 Random Noise Generator and mixing it with the noncomposite video entering the CCU. (A flat 0 -- 35 MHz mixer was designed and built for this purpose, and is described in detail in appendix A). The noncomposite signal-plus-noise was then returned to the CCU for the addition of sync and blanking waveforms, thereby avoiding the problem of adding noise to the sync and blanking signals. In order to shape the noise passband, the output from the noise generator was run through high- and low-pass passive filters prior to insertion of the noise into the mixer. Actual rms noise was monitored using a Ballantine Model 323-01 True RMS Voltmeter.

Mission input variables were manipulated largely by selecting the imagery to be presented to the TV camera. Use was made of the 35mm. imagery previously obtained from a 3000:1 terrain model located at Columbus Div. of North American Rockwell (ref. 18). The imagery contains parametric variation in groundspeed, altitude, field of view, depression angle, and shutter speed, as well as a wide variety of tactical and strategic targets in differing backgrounds.

## DISPLAY OF DYNAMIC IMAGERY

A 35mm. Norelco motion picture projector, previously modified to provide synchronization to a TV system (ref. 18) was employed to focus the 35mm. imagery directly on the vidicon photocathode. Modifications to this projector include deriving a frame pulldown sync from the flywheel of the projector, and synchronizing the TV to it. Illumination of the film frame was by a Strobex lamp, synchronized to the film pulldown and operated at 60 flashes per second, or two flashes per TV frame. Thus, the TV saw only the stabilized image of the 35mm. film while the film frame was in the projector gate.

To protect the TV monitor from "hunting" for a sync signal while the 35mm. projector was stopped, a separate sync signal was developed from 60-cycle current and used to drive the TV system. This sync generator unit also included a counter which counted frames of the 35mm. imagery as it passed through the projector gate. As the subject responded to the image on the TV monitor by closing a hand held switch, the frame number in the gate at that moment was "latched" by the counter and recorded on a printer. Simultaneously, the experimenter read the latched frame number on an LED display prior to unlatching the counter. During the latched time, the counter continued to count although the display remained frozen. Details of this sync/counter unit are given in appendix B.

## DISPLAY OF STATIC IMAGERY

Static photographs, typically 8 x 10 inches, were presented to the TV camera by mounting them in a light-controlled box, approximately 6 feet long, with photofloods illuminating the photograph. Figure 10 illustrates this arrangement, which was used for two of the experiments to be described later.

## PHOTOMETRIC MEASUREMENTS

In addition to measuring the video bandwidth of the TV chain for various configurations, it was necessary to measure the photometric characteristics of the film input and the monitor output. The reason for this is that the stimulus to

the observer is in photometric and spatial terms, and not in electric terms. Thus, if the characteristics of the display are such that the input, measured electrically, is not linearly related to the output, measured photometrically, incorrect conclusions might be drawn from using only electrical calibration of the image. As shall be seen later, this caution was well founded.

Ideally, the total transfer response of the system should be made using a sine-wave periodic intensity pattern as an optical input to the TV camera, and measuring the luminous energy as a function of position by scanning the display portion corresponding to the input pattern. However, it is very difficult to produce sinusoidally varying intensity patterns of various sizes on hard copy photographic paper, especially when it is also desirable to vary systematically the depth of the sine-wave modulation. Thus, for calibration purposes, standard 1951 USAF tri-bar patterns (square-wave intensity variation) were used. Given linearity, one can calculate the equivalent sine-wave response from such square-wave targets if necessary (ref. 19), although in the approach taken in this research it is not necessary to do so.

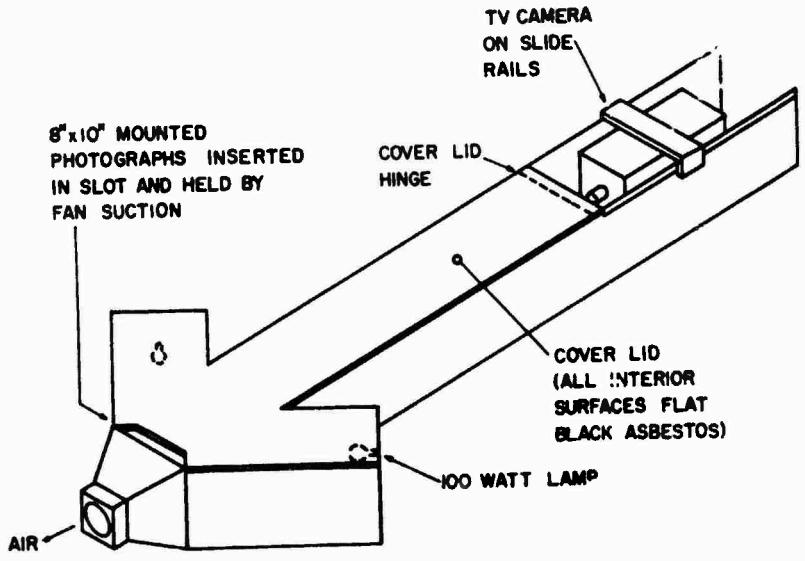



Figure 10. Tri-bar Target Exposure Apparatus.

The photometric measurements were made by using a Gamma Scientific microphotometer, equipped with a scanning eyepiece, to determine the luminance pattern or modulation of both the test target (input) and the displayed image on the TV monitor. The scanning slit in the eyepiece of the photometer subtends 25μ by 2500μ at the object plane in both cases. The output of the calibrated photometer is recorded on an X-Y plotter. Details of this procedure and the results of such measurements are given in section III of this report.

MICRODENSITOMETRY

In experiments in which the 35mm. imagery was used as the input to the television system, the imagery was previously scanned with a microdensitometer to quantify the input in equivalent luminance/spatial terms. The microdensitometer, a modification of the Gamma Scientific photometric equipment, was set to scan a 60-micron spot across the 35mm. transparency (a single frame of the ciné film) and to record the transmission through the film on an X-Y plotter. Transmission was mathematically transformed to equivalent luminance on the monitor by a calibration curve taken with a gray scale test chart made specifically for this purpose. Details of the microdensitometric experiment and results are indicated in section VI.

## SECTION III

## DETECTABILITY THRESHOLD EXPERIMENTS

One of the primary objectives of this research is tc evaluate several metrics of image quality, including the Modulation Transfer Function Area (MTFA). In order to obtain the MTFA value for any television system configuration, it is necessary to establish, analytically or experimentally, the detectability threshold curve for that set of viewing conditions and that system configuration. Two experiments and much analytical work were devoted to this end, as described in this section of the report.

### INTRODUCTION

Conventional contrast thresholds are unsuitable for MTFA calculations because they do not include the effects of the sampling process of the video raster. Coltman and Anderson (ref. 20) reported the experimental determination of some detectability thresholds as verification of a theoretical derivation. The usefulness of these curves for a thorough video evaluation of the MTFA and later applications is perhaps suggested by quoting from their paper:

> ". . .The amount of data taken was limited, and conditions of surround brightness, time interval between tests, etc., were not carefully controlled, so that the data presented here do not constitute a definitive study of this particular visual parameter. . ." p. 861

Effort was devoted to generating parametric detectability curves based upon the photographically derived MTFA (refs. 8, 9), but with little success or confidence. Fundamental differences between photographic imagery and line-scan electronic displays lie in the areas of noise (static for photographic, dynamic for CRTs); raster inferference in one dimension (not present in photographs); possible differences in gamma, both in slope and overall curve linearity; mean viewing luminance levels; dynamic display range; and shape of the noise spectrum (noise passband, such as white vs. narrow band). Thus, an experimental determination of line-scan detectability thresholds was indicated.

Two experiments were conducted to determine empirical detectability thresholds for a variety of video system conditions. The first experiment, essentially a pilot study, resulted in the threshold detectability curves to be reported in section IV. This first study, while it generated the single set of five curves needed for the calculations in section IV, did not produce any data of generalizable utility, and is therefore of limited value. In the interest of brevity, and because the techniques for the two threshold experiments are essentially equivalent, the first experiment will not be discussed at length, nor will the results be given here. Rather, emphasis will be placed upon the second, more inclusive, experiment.

## DESIGN OF THE SECOND EXPERIMENT

Figure 11 summarizes the specific experimental design. Three TV system configurations, each with a different MTF, were used. The three different MTFs were obtained by operating the variable-parameter TV at 32 MHz with a 1225 lines-per-frame rate, at 16 MHz bandwidth and 945 lines, and at 8 MHz bandwidth, 525 lines. The square-wave photometric responses (tri-bar modulation output/tri-bar modulation input) corresponding to these three MTFs are given in figure 12. Three or four different noise passbands at each bandwidth/line rate combination were included, as shown in figure 11. The targets were a series of 8 x 10-inch photographs of a single 1951 USAF tri-bar pattern with darker bars against a white background. The targets were made in seven spatial frequencies with eight modulations at each spatial frequency and were displayed with the major axis of the bars perpendicular to the TV raster lines. The raster was oriented vertically, as viewed by the subject.

Subjects were three male and two female paid students having a minimum of 20/22 binocular and 20/25 monocular near and far visual acuity without correction, and without any visual anomalies as tested with the Bausch & Lomb Ortho-rater. Following approximately 100 practice trials, each subject received six trials at each combination of spatial frequency, modulation, noise passband, and bandwidth/line rate, for a total across all subjects of 18,480 data points. The subject, monitor, noise generator, noise filters, photometer and the first

experimenter were in one room. In an adjoining room were the camera, CCU, targets, and the second experimenter. An intercom connected the two experimenters. Both rooms were climate controlled.

Each subject was seated in a variable position dental chair, and instructed to lean against a headrest so that the eye-to-CRT distance was 40 in. and the line of regard was normal to the center of the CRT. The noise generator was located so that the noise level potentiometer could be adjusted by the subject while the rms noise level meter was visible only to the experimenter. The experimental room was dark except for the monitor and a small lamp for the experimenter. No light from this small lamp fell in the subjects' field of view.

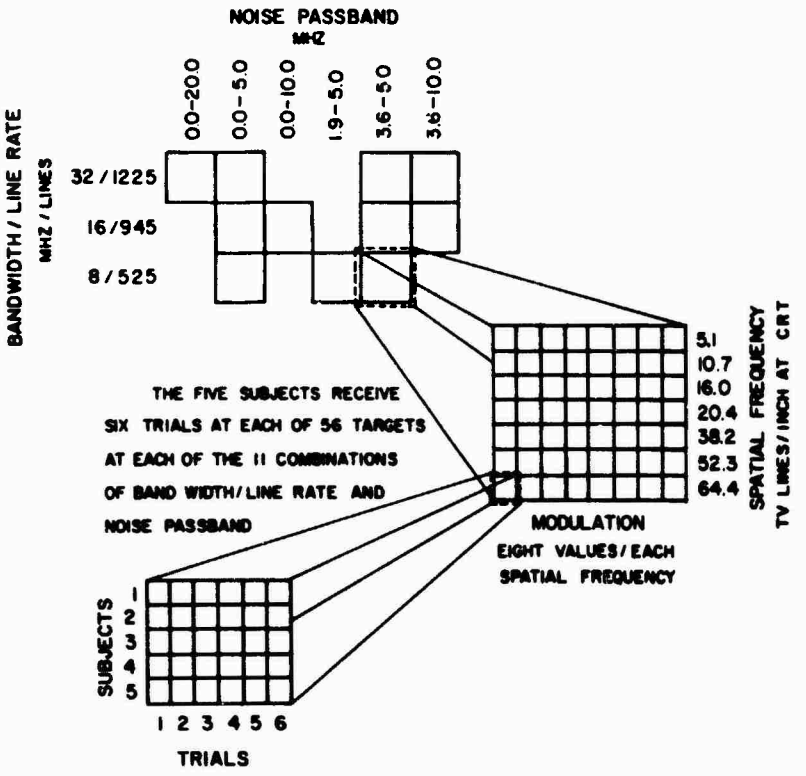



Figure 11. Threshold Detectability Experimental Design. Note that 1 TV line/inch at the CRT is equal to 0.349 cycles per degree for the given viewing conditions.

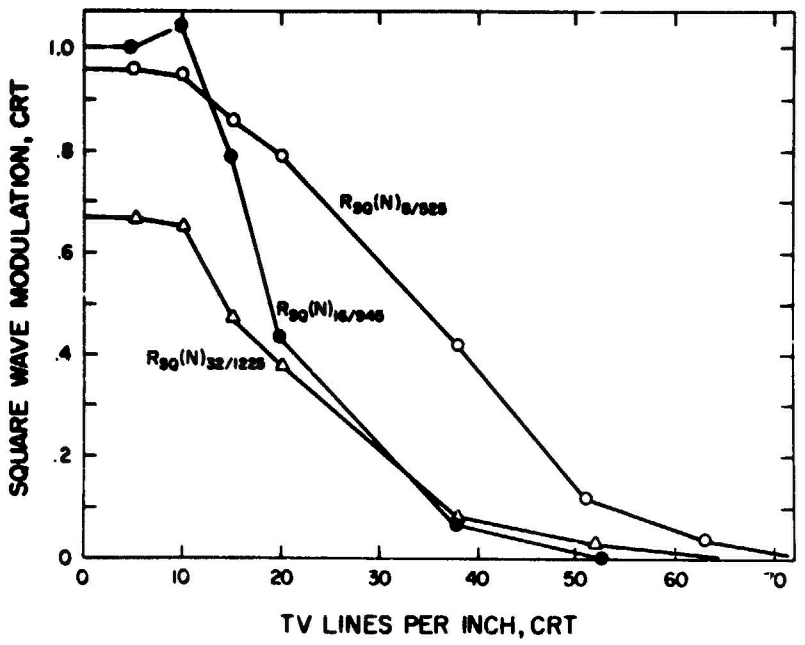



Figure 12. System Square Wave Response Functions.

## PROCEDURE

The normal daily experimental procedure began with one hour of warm-up for all of the electronics prior to the start of data collection. At the end of this warm-up, the system was checked for calibration using a 10-step gray scale target in front of the camera. Overall video and blanking levels at the CCU were checked and adjusted. Composite video from the CCU was monitored on an oscilloscope and any necessary CCU adjustments were made to return the video levels for each gray bar to predetermined set values. With the electrical input to the monitor thus standardized, the image on the monitor of the gray scale target was viewed with the telephotometer. The luminance of three particular gray bars, one near white, the second middle gray, and a third near black, was measured, and the contrast and brightness controls of the monitor were adjusted to bring the luminance of these bars within certain tolerances (approximately 0.2 ft-Lamberts). This calibration procedure was performed at the beginning of each experimental session and repeated every hour during the session. As an example of the importance of these procedures, the luminance of the nearly white bar was adjusted during calibration to be between 18.0 and 18.5 foot-Lamberts. After one hour of operation, the luminance would drift 1 or 2 foot-Lamberts. There is no known evidence to suggest that this amount of drift is peculiar to this particular TV system.

Note, in figure 12, that the $R_{sq}(N)$ values are generally poorest for the 32 MHz 1225 line system, and best for the 8 MHz, 525 line system. These differences are perhaps due to the foregoing standarized set up procedure, and may not represent maximum image quality as viewed subjectively or as measured by other (e.g., electrical) image quality metrics. Nonetheless, these curves accurately depict the visual stimulus to the observer, are therefore the most valid representation of the physical displayed image, and, as shall be seen in subsequent sections of this report, predict accurately the performance of observers. Thus, the nature by which such $R_{sq}(N)$ values are generated is unimportant compared to the validity of measuring such values and their ultimate use in defining image quality.

After calibration, a subject was seated and the seat height was adjusted so that the subject's eyes were in proper position and the subject was comfortable. The first experimenter requested a target at random. The second experimenter placed the requested target before the camera. The subject increased the noise level until he could no longer determine that there were three separate bars. The first experimenter recorded this noise level, increased the noise level until well past the point where the target was not visible, and told the subject to proceed. The subject then reduced the noise level until he could just determine that there were three separate bars. This noise level was then recorded, completing the pair of trials. The subjects' criterion was not that of a detection task or a recognition task; the criterion was simply the existence or non-existence of three, separate bars.

In theory, since the ascending trial gave a noise level slightly higher than the actual threshold noise level and the descending trial gave a noise level slightly below the threshold, the mean of these two trials is taken as the threshold. With practice, a pair of trials was completed in about 12 seconds. Subjects worked for periods of 20 to 25 minutes and were then given a 5-to-10-minute break. Three of these periods plus the calibration procedures filled a 2-hour experimental session. Each of the five subjects worked each day during the experiment and for no more than 2 hours. This work was demanding, and any more than 2 hours per day per subject caused subject fatigue and erratic performance. Each subject worked approximately 14 hours during the experiment.

## RESULTS

Means of each combination of spatial frequency, modulation, noise passband, and bandwidth/line rate were calculated and plotted. To describe quantitatively the results, simple linear regressions were calculated with modulation predicting threshold noise level at each spatial frequency and at each of the eleven bandwidth/line rate-noise passband combinations. Multiple linear regressions of spatial frequency and modulation predicting threshold noise level were also obtained at each bandwidth/line rate-noise passband combination.

Figures 13 through 16 show the relationshop between square-wave modulation and the mean threshold input noise level, given in rms millivolts, for the four noise passbands used in the 32 MHz bandwidth/1225 line system.

Figure 13 is for the 0 - 20 MHz noise passband. The relationship among the parameters is as expected. At any spatial frequency, as modulation increases, the threshold rms noise also increases. As spatial frequency increases (target size diminishes), the threshold noise declines. For the largest target, having a spatial frequency equal to 5.1 TV lines/inch, the threshold curve is nonlinear. In the segment where modulation is less than about 0.3, any increase in modulation is matched by an increase in the noise threshold. Beyond this segment, an increase in modulation allows little or no increase in noise. A plateau of sorts has been reached. For the largest targets with higher contrast, most of the subjects commented that the whole target disappeared at the point in increasing noise where they could no longer determine the existence of three separate bars. In other words, the target was totally obliterated at the noise threshold. For smaller targets, the target remained visible as a "smudge" as noise was increased above the threshold, even though three separate bars could not be seen.

In figure 13, as well as in those following, the maximum modulation plotted at each spatial frequency declines with increasing spatial frequency, illustrating the rolloff of the square-wave response as spatial frequency increases.

Although targets of 7 spatial frequencies were used at each of the 11 system combinations, only the lower spatial frequencies are shown in these graphs. Even with a fair range of modulation on the photographic prints of the higher spatial frequencies, the displayed modulation range is quite small due to the TV system square-wave response rolloff, thus yielding mean threshold curves of only two or three points. For clarity, these curves of limited usefulness were not plotted.

The detectability threshold curves for the 0 - 5 MHz noise passband given in figure 14 show the same ordering of points and general shape as those in

figure 13. These threshold noise levels are only half as great as those at the 0 - 20 MHz passband. Since the 0 - 5 MHz band is contained in both of these noise passbands, it is concluded that the 0 - 5 MHz noise band includes the most detrimental noise frequencies. This conclusion is subjectively reasonable, since the lower frequencies will cause the larger "snow flakes". The energy expended in the 5 - 20 MHz noise frequencies apparently has much less effect on target detectability.

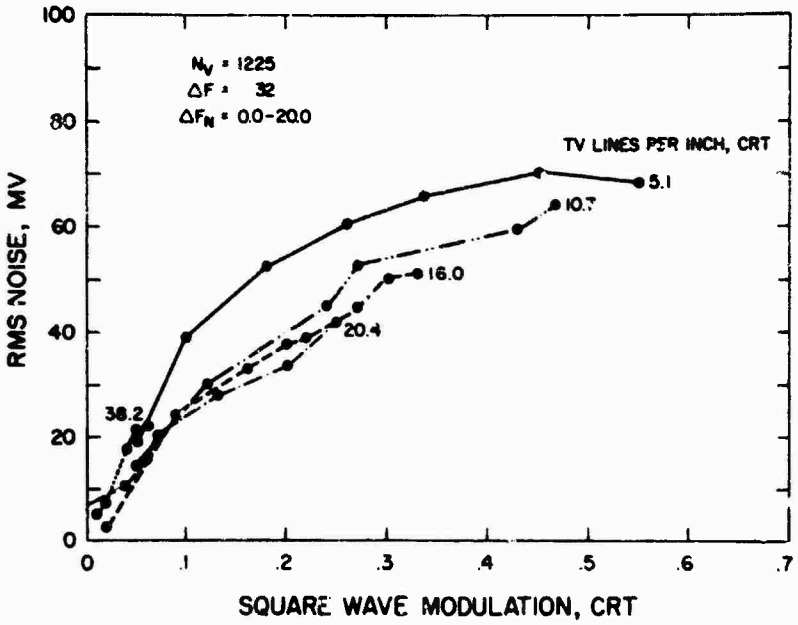



Figure 13. Empirical Noise Threshold Data.

The detectability thresholds for the 3.6 - 5 MHz noise passband given in figure 15 are slightly lower than those of the 0 - 20 MHz noise passband. If the noise in the 3.6 - 5 MHz region were an important contributor to the impairment of tri-bar detection, then the rms noise voltages required to reach threshold conditions would be substantially less than that rms noise required at the 0 - 20 MHz noise passband. But such is not the case. More importantly, the threshold rms noise voltage for any target is much higher at the 3.6 - 5 MHz noise passband than at the 0 - 5 MHz noise passband, indicating the important noise frequencies are between 0 and 3.6 MHz.

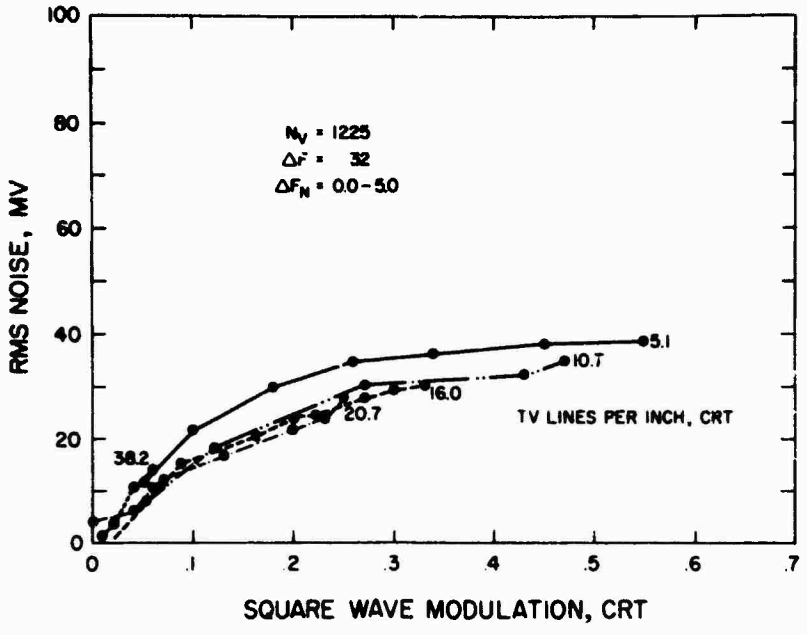



Figure 14. Empirical Noise Threshold Data.

Because noise in the 3.6 - 5 MHz passband was relatively ineffectual in degrading targets, then noise energy between 3.6 and 10 MHz is even more wasted. These thresholds, shown in figure 16, are the highest in the 32 MHz bandwidth/1225 line group.

The detectability thresholds for the 16 MHz/945 line and 8 MHz/525 line systems are presented in figures 17 through 23. The ordering and shape of the curves are similar to those of the 32 MHz/1225 line systems. The detectability thresholds in the 8 MHz/525 line system with a 1.9 - 5 MHz noise passband are considerably higher than those for the 0 - 5 MHz noise passband at the same bandwidth/line rate. This result suggests that the most detrimental noise frequencies are below about 2 MHz.

At the 525 line rate, on a 10 x 14 in. monitor, 2 MHz converts to about 18 TV lines/inch at the monitor. As can be seen in figure 12 at this point the square-wave response for the 8 MHz, 525 line system has started to decline. One might conclude that the reduced contribution of noise frequencies greater than 2 MHz is due to the system response rolloff above 2 MHz; that is, frequencies greater than 2 MHz have less amplitude due to the monitor square-wave response.

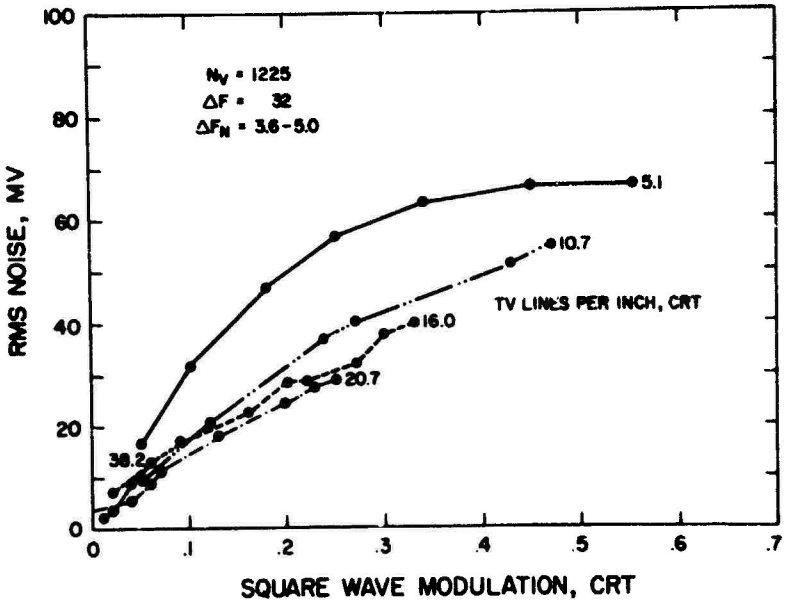



Figure 15. Empirical Noise Threshold Data.

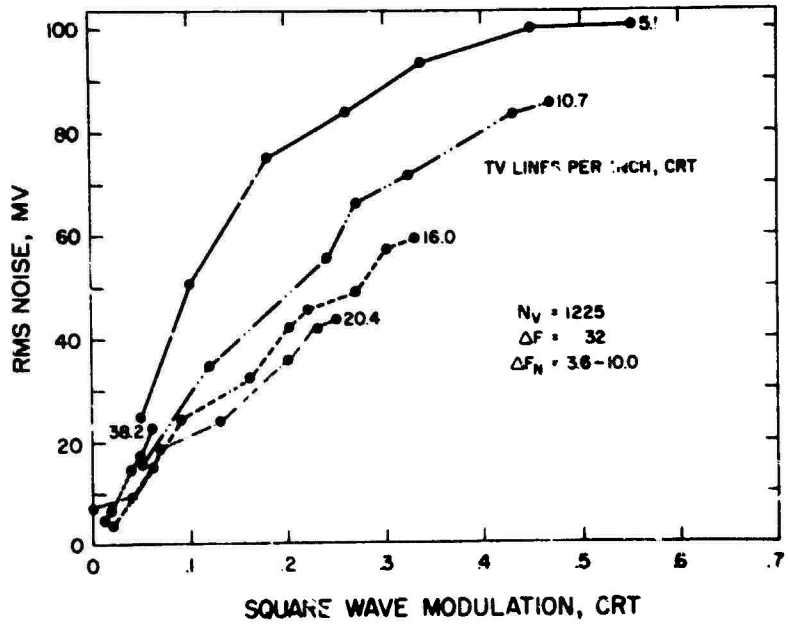



Figure 16. Empirical Noise Threshold Data.

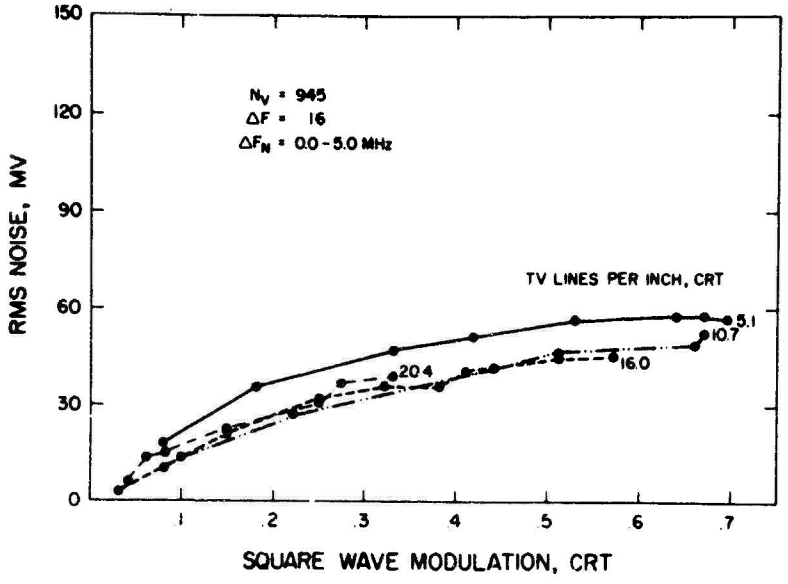



Figure 17. Empirical Noise Threshold Data.

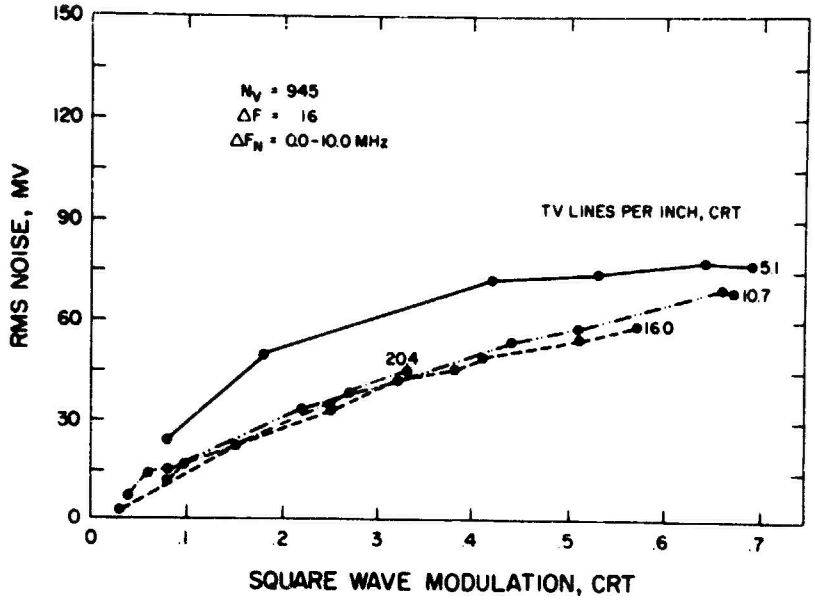



Figure 18. Empirical Noise Threshold Data.

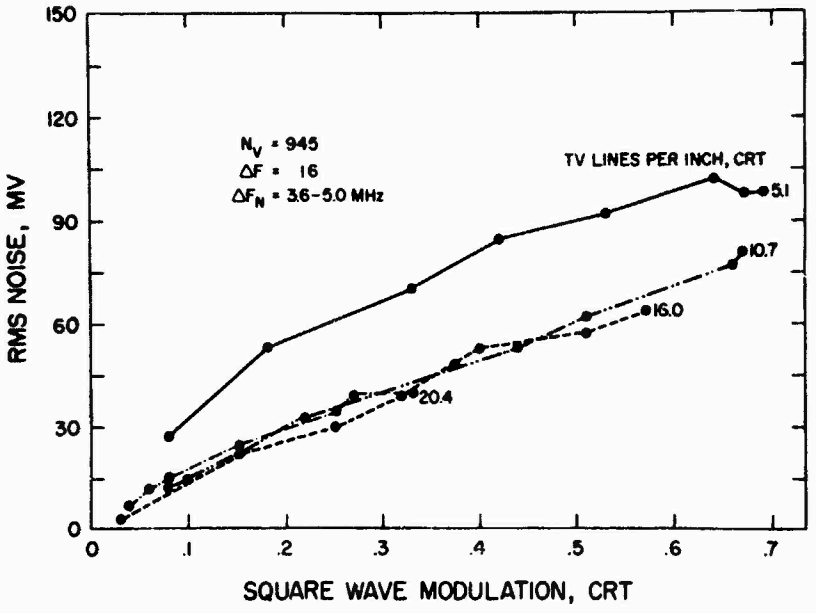



Figure 19. Empirical Noise Threshold Data.

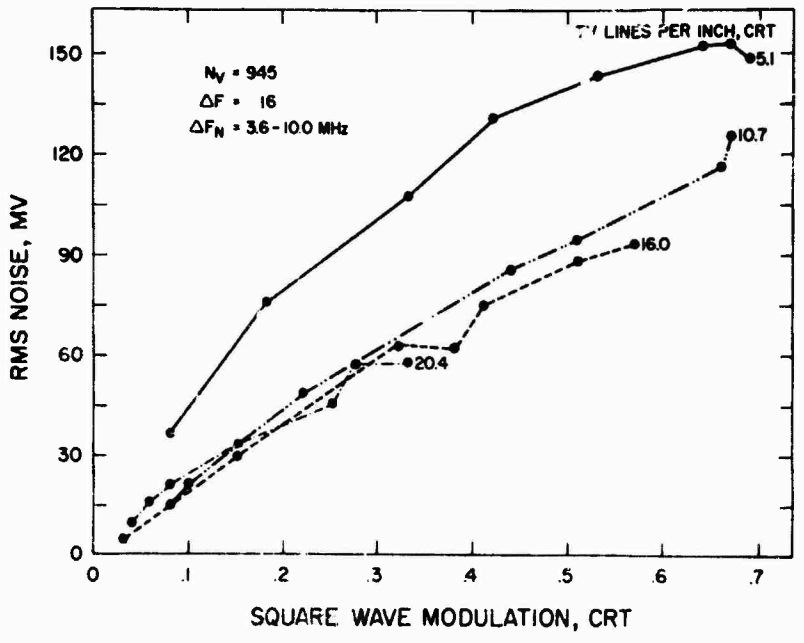



Figure 20. Empirical Noise Threshold Data.

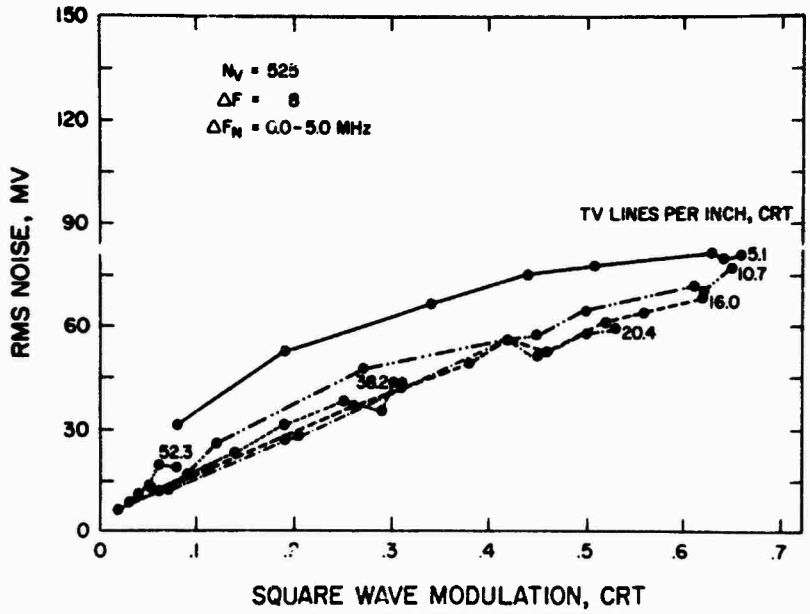



Figure 21. Empirical Noise Threshold Data.

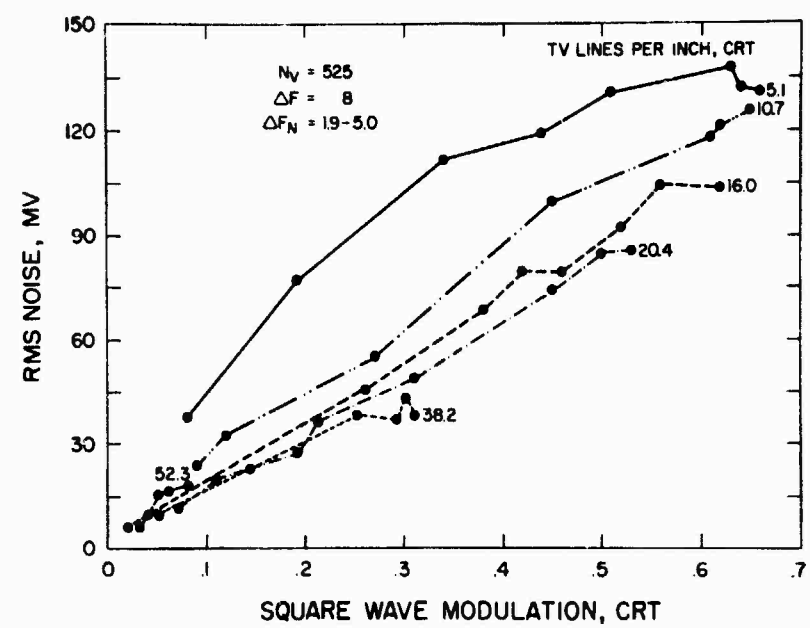



Figure 22. Empirical Noise Threshold Data.

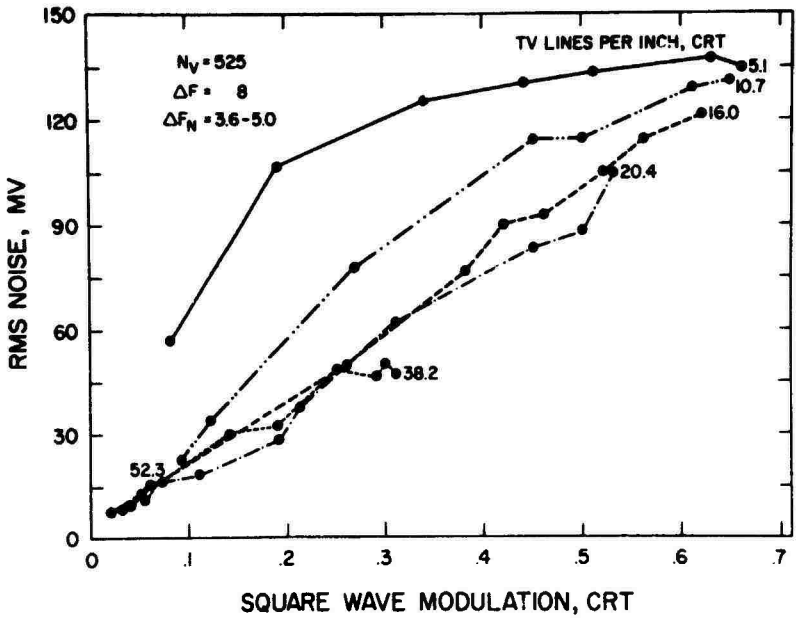



Figure 23. Empirical Noise Threshold Data.

This is only a partial explanation of the importance of low frequency noise, however. At the 1225 line rate, the 3.6 - 5.0 MHz noise passband falls between 15 and 20 TV lines/inch spatial frequency at the monitor. The square-wave response for the 32 MHz/1225 line system is relatively flat to about 10 TV lines/inch at the display although the square-wave modulation response is only about .67. So, the elevation of the noise thresholds for the 3.6 - 5.0 and 3.6 - 10.0 noise passbands is due to both the differential rolloff of the higher-frequency noise by the system square-wave response and to the inherent detriment of lower frequency noise, i.e., noise less than about 2 MHz.

## DISCUSSION

### Regression

Simple and multiple linear regressions were originally applied to the data for several reasons. First, a simple algebraic description of the results was sought to convert the data to a form useable in MTFA calculations. The original data were in the form of threshold noise level as a continuous variable dependent on discrete values of spatial frequency and modulation for each noise passband-bandwidth/line rate combination. For each combination, an applied linear regression is of the form:

$$a(SF) + b(M) + c = \sigma_N \qquad (8)$$

where SF is the spatial frequency in TV lines/inch, M is modulation, $\sigma_N$ is the threshold noise level in rms millivolts, and a, b, c are constants of the regression.

Also, the MTFA requires threshold curves relating spatial frequency and modulation with noise level as a discrete parameter. That is, the required algebraic equation is of the form:

$$M = a(SF) + c \qquad (9)$$

with $\sigma_N$ equal to some particular constant. The regression equation (8) can be solved for modulation (M) to satisfy the MTFA requirement. Figure 24 plots, as an example, the multiple linear regression for the 32 MHz/1225 line, 0 - 20 MHz noise passband data.

A good question is the appropriateness of the linear, rather than a nonlinear, regression. Linear regressions were originally fit for speed and simplicity. As shown in the summary of the multiple linear regressions, table 2, the minimum correlation coefficient is .90, which can be interpreted as meaning that at least 81% (100 x $.90^2$) of the variance is predicted by the linear equation. As expected, all F tests for regression effects proved highly significant. This result does not indicate the linear model is the best-fitting one, however. An F test for lack of fit applied to the simple linear regression for the 32 MHz/1225 line, 0 - 20 MHz noise passband condition, and spatial frequency equal to 5.1 TV lines/inch was significant ($p < .001$), indicating that the linear model was incorrect. In a rash attempt to find a better model, a stepwise multiple regression was applied to 27 different transformations, (e.g., reciprocal, logarithm c, arcsin) of spatial frequency and modulation using the 32 MHz/1225 line, 0 - 20 MHz noise passband results. Using an extremely liberal F = .01 for inclusion and F = .005 for deletion, the regression program stopped after 12 steps at a multiple correlation coefficient of .94, indicating that no further increase of the coefficient was possible by either adding or deleting any of the 27 transformations. Since the multiple R for the linear model is .92 (table 2), the improvement in percentage of the variance predicted by the regression is 3% [$(.94)^2 = .88$, $(.92)^2 = .85$]. Such a small improvement does not seem to warrant the use of an equation of twelve off-beat variables to predict threshold noise. This analysis suggests that 12% of the variance among means is truly random and unpredictable. On the assumption that the models underlying the other ten noise passband-bandwidth/line rate combinations are of similar form, the linear equations can be used with the realization that they are not perfect, but rather useful, very close approximations.

Spot-Size Limitation

The order of square-wave response functions for these three systems (figure 12) is reversed from prior expectations, perhaps due to the previously discussed standardized set-up procedure, but also due to the spot-size limitation of the particular monitor used here. Although spot-size measurements were not made, spot size increased very likely with the faster writing speed of the higher line rates. This offers a possible explanation for the increase in threshold noise

level with decreasing bandwidth/line rate. If the spot sizes for the 8 MHz/525 line system and for the 32 MHz/1225 line system were equal, then the impairment due to a given amount of inserted noise would be greater in the 32 MHz/1225 line system because the greater writing speed results in a larger "snow flake" for any positive noise pulse. This explanation will be evaluated during the second phase of this research.

The reduced square-wave response with increased line rate may also be due to raster instability. It has been informally reported by others that this particular model monitor suffers from considerable frame-to-frame raster instability, which would account for its slow-scanned photometric response being much poorer than its electrical response.

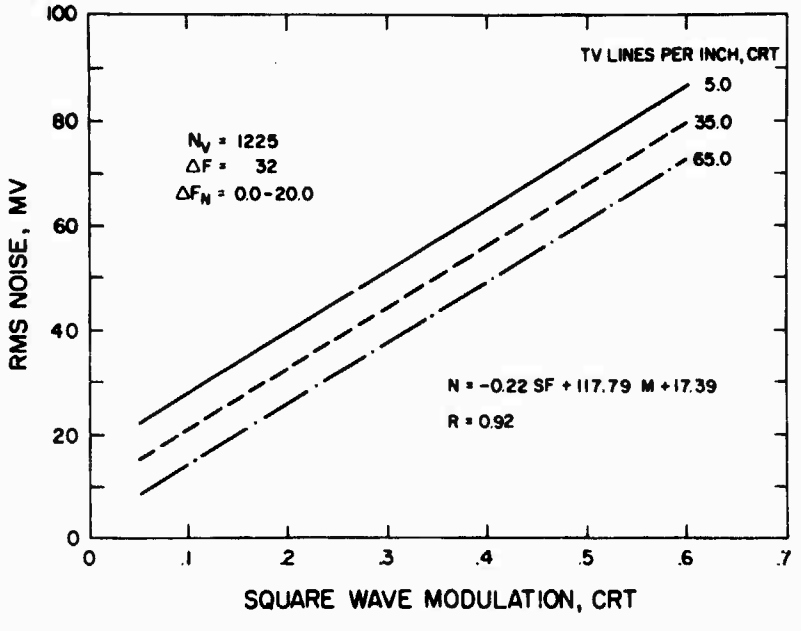



Figure 24. Threshold Detectability Multiple Linear Regression.

Table 2. MULTIPLE LINEAR REGRESSION EQUATIONS FOR NOISE THRESHOLDS

| $\Delta F/N_V$ | $\Delta F_N$ | R | N = ____ | SF + _____ | M + _____ |
|---|---|---|---|---|---|
| 32 MHz/1225 lines | 0.0 - 20.0 MHz | .92 | -.22 | 117.79 | 17.39 |
| | 0.0 - 5.0 MHz | .94 | -.16 | 63.61 | 11.41 |
| | 3.6 - 5.0 MHz | .93 | -.16 | 112.89 | 10.04 |
| | 3.6 - 10.0 MHz | .93 | -.24 | 171.73 | 15.73 |
| 16 MHz/945 lines | 0.0 - 5.0 MHz | .95 | -.24 | 68.44 | 14.65 |
| | 0.0 - 10.0 MHz | .90 | -.27 | 92.47 | 16.64 |
| | 3.6 - 5.0 MHz | .94 | -.21 | 118.95 | 11.85 |
| | 3.6 - 10.0 MHz | .91 | -.29 | 185.01 | 14.99 |
| 8 MHz/525 lines | 0.0 - 5.0 MHz | .91 | -.28 | 95.01 | 20.44 |
| | 1.9 - 5.0 MHz | .91 | -.42 | 162.75 | 22.07 |
| | 3.6 - 5.0 MHz | .91 | -.58 | 168.75 | 31.30 |

SYSTEM PHOTOMETRY

Besides describing quantitatively the target modulation and spatial frequency put into the system and the displayed modulation and spatial frequency at the monitor, the relationship between the electrical noise inserted into the system and the displayed luminous noise was also sought.

Input spatial frequency was physically measured on the photographic prints. The modulation of the targets on these prints was measured using a microphotometer with a 25μ by 2500μ scanning slit eyepiece. Spatial frequencies of the targets at the display were calculated from the input spatial frequency and the system magnification. A sample of targets was measured at the display to verify these calculations. Target modulation at the display was found using the same microphotometer and scanning eyepiece as used in the input modulation measurement of the photographic prints. This displayed modulation was measured at each different bandwidth/line rate combination since the system response $R_{sq}(N)$ changed.

Fortunately, the output rms voltmeter of the noise generator proved to be accurate as checked by the true rms voltmeter, and its readings were used for the input noise level. Corrections for attenuation of the noise passband were applied in the data analysis.

The size of the scanning spot is also an important parameter of any line-scan display. An attempt to measure spot size was made using a high efficiency microphotometer with a double slit aperture (0.003 x 0.400 inch with 0.150 inch spacing at the CRT). Theoretically, the spot passing the two slits should give two peaks on an oscilloscope displaying the output of the photomultiplier. From these two peaks, the spot size can be found. At these line rates, and for this particular CRT, the persistence of the P4 phosphor was so long relative to the speed of the spot that only one peak was obtained on the oscilloscope for each passing of the spot across the two slits.

This photomultiplier tube output presented a possible measure of the variation of the spot luminance due to inserted noise, however. The photograph of one of these oscilloscope images is given in figure 25. The output of the photometer was taken directly from the dynode of the photomultiplier tube. Evidence to the absence of saturation of the photomultiplier tube is given by the lack of limiting seen in the peaks in figure 25. With the photomultiplier tube operating in its linear region, and the results displayed on a calibrated Tektronix 7403N oscilloscope, the peak height is linearly proportional to the integrated luminance of the imaged area as the spot passes the aperture. The standard deviation of the peak heights, when converted to ft-Lamberts, is then proportional to the rms luminance of the spot and can be compared with the rms noise voltage inserted.

The results are given in figures 26 through 36. These graphs show, for each bandwidth/line rate-noise passband combination of the threshold experiment, the relationship between the inserted noise in rms volts and the displayed noise in arbitrary units of rms luminance. These measurements were made against both large white and large black image areas. The range of inserted noise at each system combination was determined by the approximate range of the detectability thresholds at these system combinations.

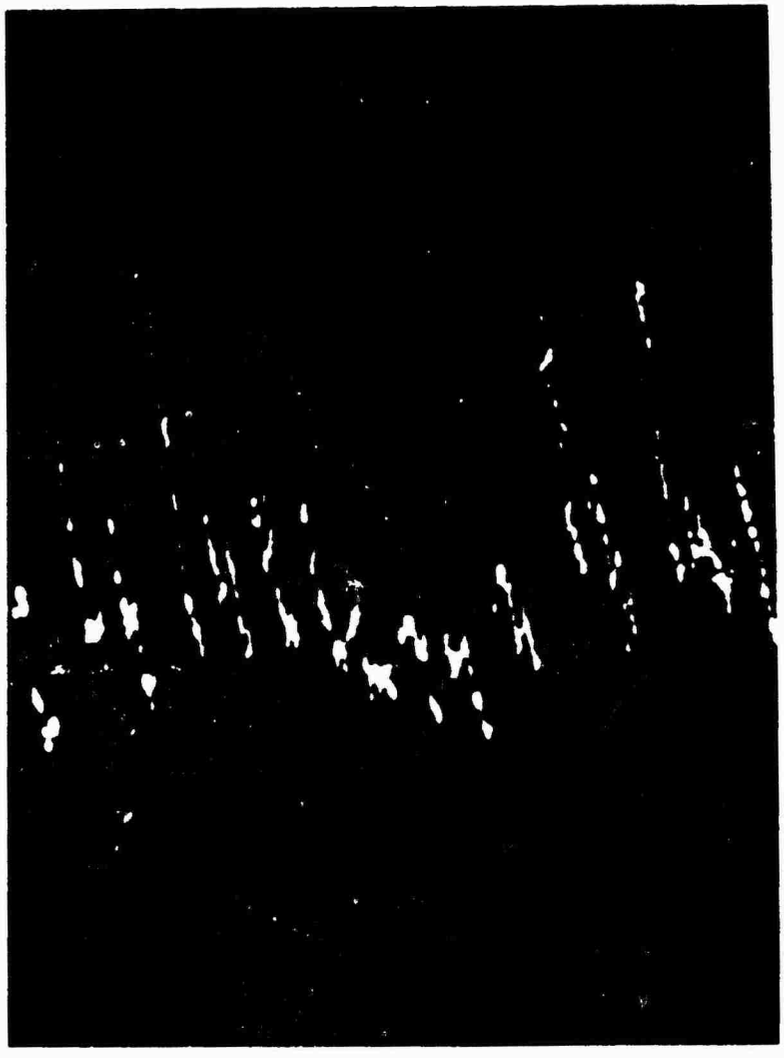


Figure 25. Photograph of Oscilloscope Trace of Noise Measurement at CRT

Before further discussion of figures 26 to 36, it should be pointed out that there are data missing in figures 30, 31, and 32. Some points near zero-inserted noise are not plotted for simplicity. Each point in the 32 MHz/1225 line system combinations is based on about 36 measured peaks, each point in the 16 MHz/945 line plots is based on about 28 peaks, and about 16 peaks were measured for each plotted point of the 8 MHz/525 line data.

An inspection of these graphs indicated that application of any linear relation is inappropriate. In most cases, the curves are made up of two linear portions. For instance, in figure 29 (32 MHz/1225 line, 3.6 - 10.0 MHz noise passband), increasing inserted noise up to about 0.025 rms volts results in a significant

increase in the rms luminance. After this point, increasing inserted noise results in a slight decline in luminance variation.

Another consideration in this interpretation of the results is that the mean luminance, and thus the modulation, is not constant with increasing amounts of inserted noise. As the inserted noise increases, the mean luminance of a white area remains fairly stable but the mean luminance of a black area increases at the same rate as the rms luminance to about the point of peak rms luminance, at which time both the rms luminance and the mean luminance stabilize.

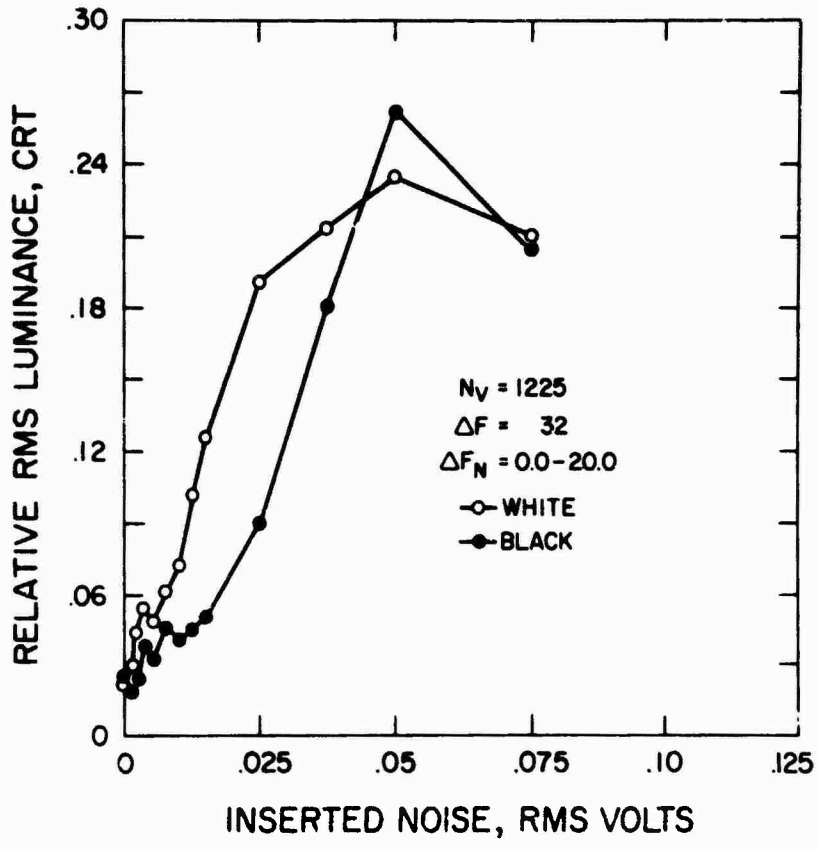



Figure 26. Photometric Noise Output.

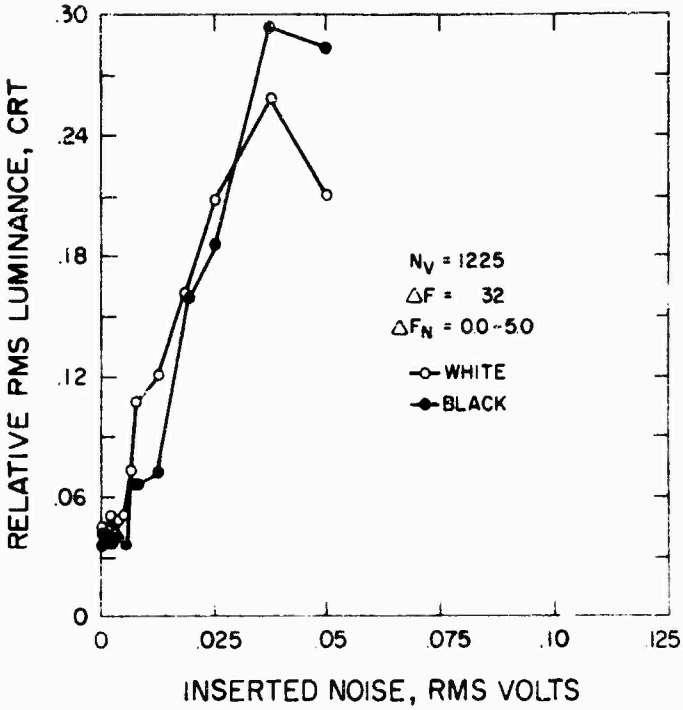



**Figure 27. Photometric Noise Output.**

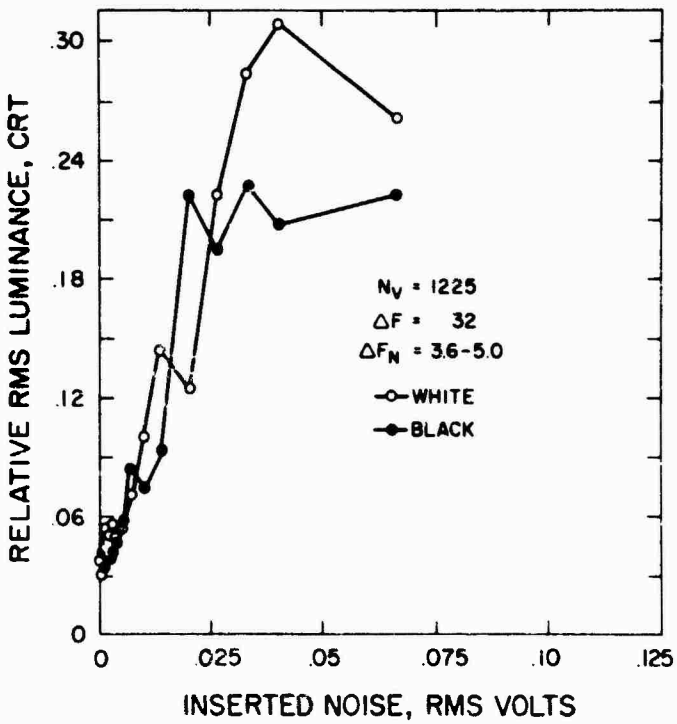



**Figure 28. Photometric Noise Output.**

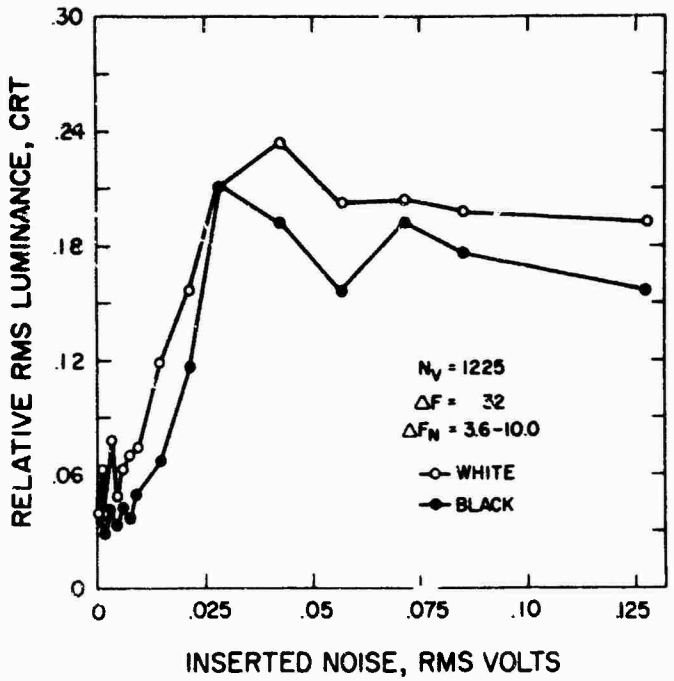



Figure 29. Photometric Noise Output.

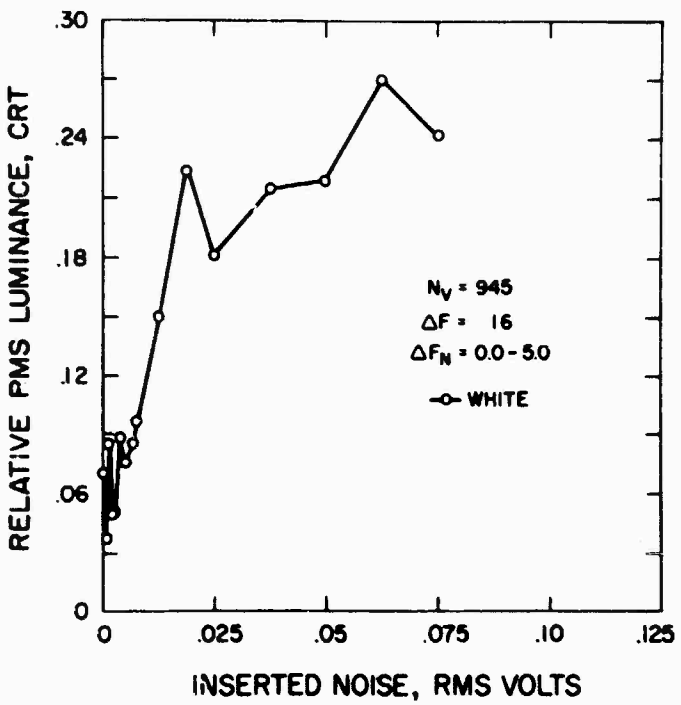



Figure 30. Photometric Noise Output.

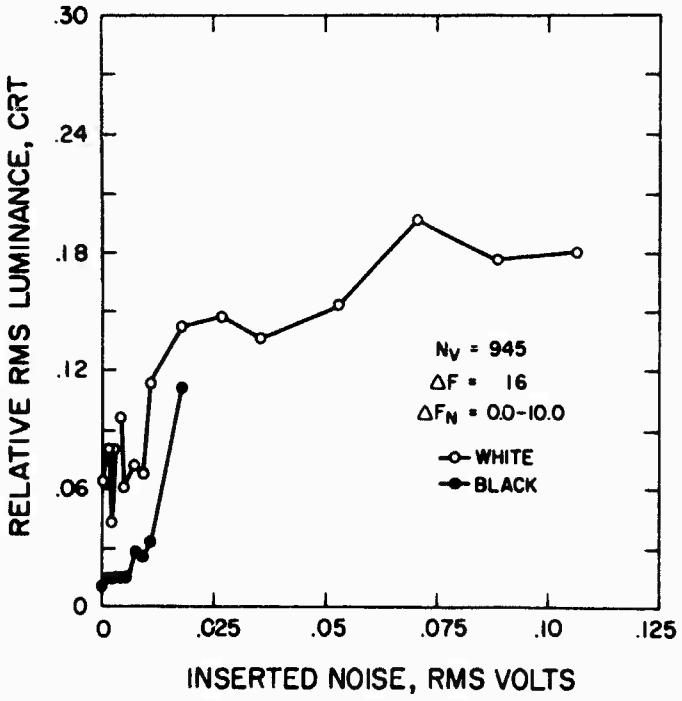



Figure 31. Photometric Noise Output.

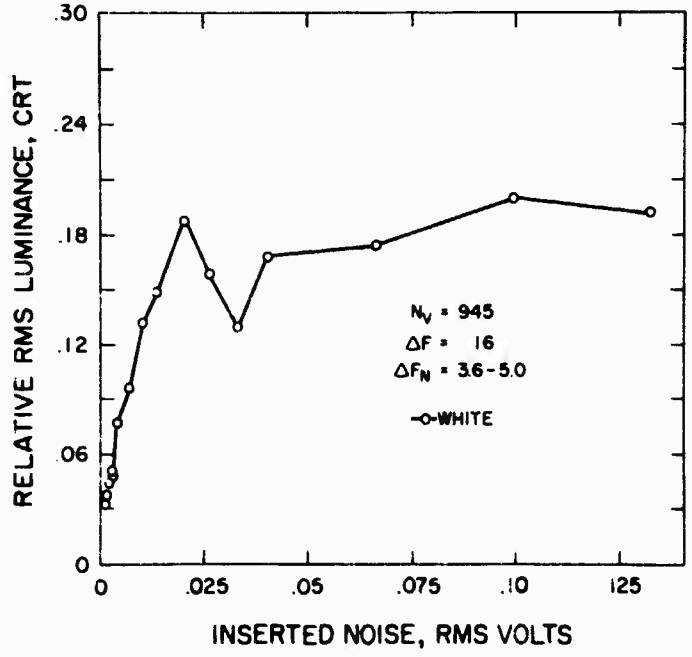



Figure 32. Photometric Noise Output

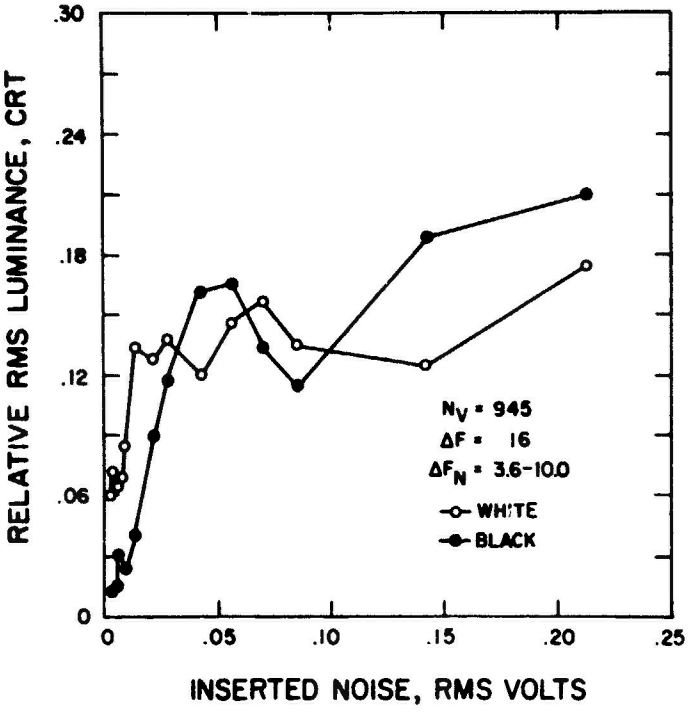



Figure 33. Photometric Noise Output.

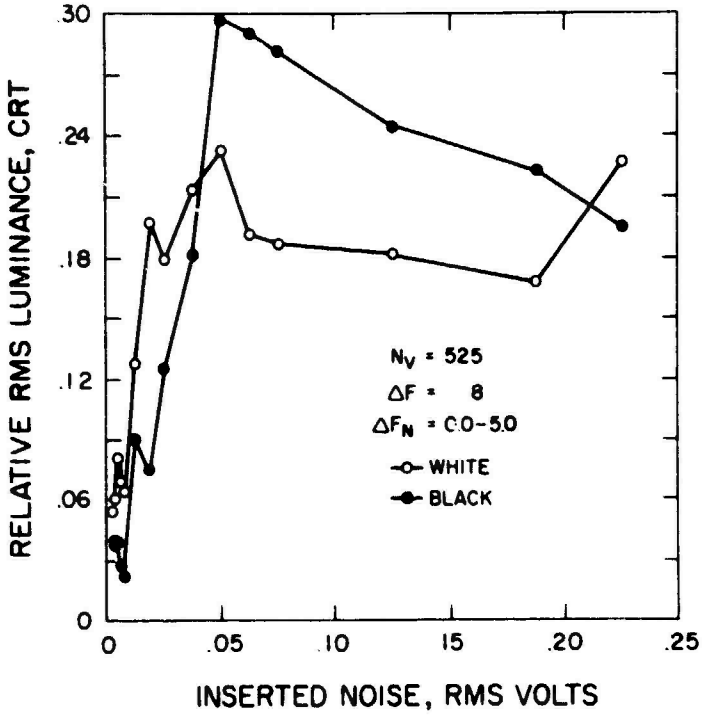



Figure 34. Photometric Noise Output.

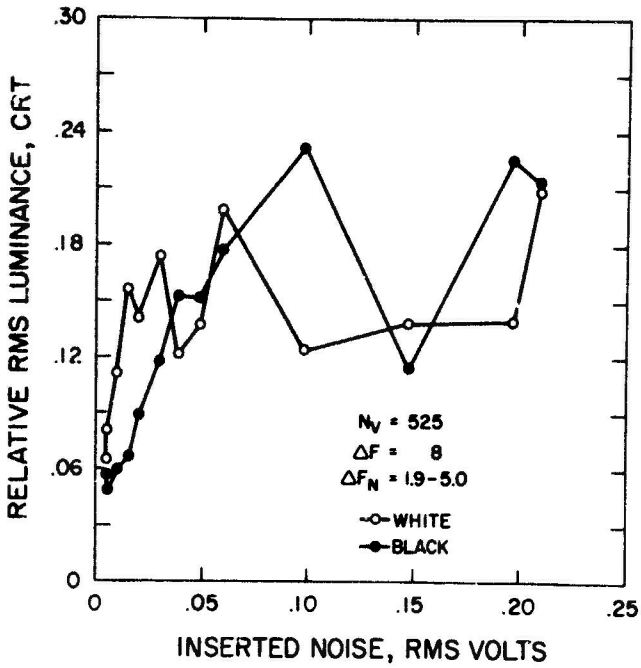



Figure 35. Photometric Noise Output.

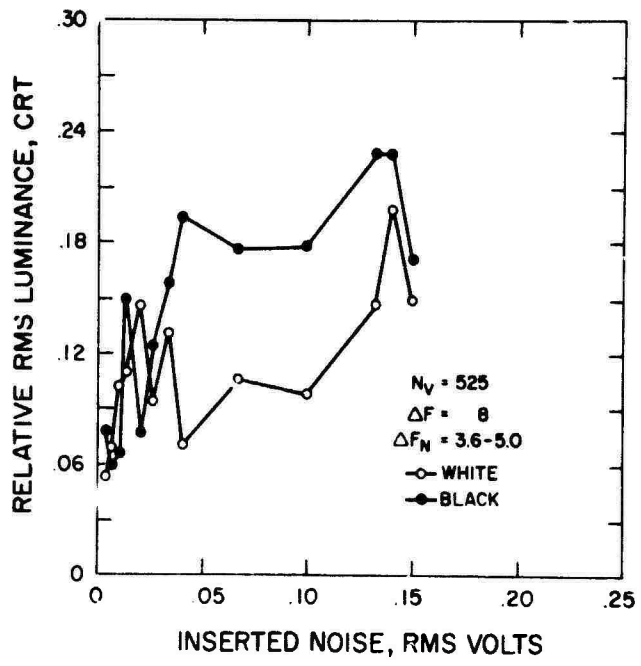



Figure 36. Photometric Noise Output.

The relationships among rms luminance, mean luminance, detectability thresholds, and target modulation certainly need further study. There may also be an interaction between the stability of the mean luminance and rms luminance at higher noise levels, and the spot-size limited nature of this (or any other) monitor.

Because the primary purpose of this display photometry was the specification of display parameter values involved in finding the detectability thresholds, the photometric noise data are limited and should certainly not be considered definitive.

## CONCLUSION

As expected, at any system and noise-passband combination, an increase in modulation at a particular spatial frequency produced an increase in the noise threshold; and at any particular modulation, an increase in spatial frequency brought about a decrease in the detectability threshold.

The ordering of the square-wave response $R_{sq}(N)$ curves caused a corresponding ordering of the detectability thresholds. The lowest overall square-wave response was that of the 32 MHz/1225 line system and the lowest overall noise detectability thresholds were also obtained for the 32 MHz/1225 line system.

Noise passband is very important in setting the detectability thresholds. The data from the different noise passbands indicate that lower frequency noise, less then 2 MHz, caused the greatest increase in the noise thresholds.

Finally, the photometry of displayed noise reported herein was done only to quantify the stimuli used to find the detectability threshold. There is a definite further need to find the relationship between electrical noise and actual displayed noise. Research which assumes linearity between inserted rms electrical and displayed rms luminous energy noise may be in error, and should be reevaluated.

## SECTION IV

## DYNAMIC TARGET RECOGNITION EXPERIMENTS

Two experiments were conducted to evaluate the alternate measures of TV image quality for dynamic air-to-ground target acquisition. The first experiment was designed essentially to check out the experimental procedures and measurement techniques for the video system using motion picture imagery. Although the operator performance data from this first experiment were more or less as predicted for the combinations of video bandwidth, line rate, and noise level, serious problems were encountered in measuring the photometric qualities of the system, which made such comparisons totally unreliable. Thus, to avoid discussing the fairly obvious and unimportant results, this first experiment will be deleted from this report.

The second dynamic imagery experiment, described below, was designed to compare target acquisition performance for five different noise levels with a constant line rate/video bandwidth system, and to relate observer performance to the alternate measures of image quality.

### EXPERIMENTAL DESIGN

The video system was set at a line rate of 945 lines/frame and a video bandwidth of 16 MHz (-3dB). Films were selected from the 35 mm. library which simulated flight at a ground speed of 500 ft./sec. and an altitude of 10,000 ft. The taking camera field of view was 41 degrees horizontal by 52.9 degrees vertical, with the boresight depression angle set at 45 degrees. The film frame was underscanned by the 3:4 aspect ratio television camera, such that the field of view as presented on the TV monitor was approximately 40 degrees vertical by 30 degrees horizontal, with a boresight depression angle of 45 degrees.

Five noise levels were obtained by adjusting the noise input to the video mixer. The noise passband was 20 Hz to 16 MHz; that is, the noise passband matched the video passband. The five noise levels, their equivalent signal-to-noise ratios (assuming a 100% contrast target input) in the video, and the signal-to-noise levels in decibels are given in table 3.

Table 3. CONDITIONS STUDIED IN THE SECOND DYNAMIC IMAGERY EXPERIMENT

| Noise Inserted, $\sigma_N$ | Highlight S/N | Highlight S/N, dB |
|---|---|---|
| 0 | 32 | 30 |
| .006 V. | 13.84 | 20.0 |
| .013 V. | 6.66 | 16.4 |
| .020 V. | 4.50 | 13.0 |
| .027 V. | 3.34 | 10.4 |

Eleven subjects, eight males and three females, were randomly assigned to each of the five different noise levels. Each subject was checked for normal vision, using the Bausch & Lomb Ortho-Rater and requiring a 20/20 near and far acuity criterion for both eyes and no worse than 20/30 for each eye, independently. Upon arriving at the laboratory, each subject was asked to study the target photo book containing the 25 targets with their backgrounds masked out. The targets, indicated by an arbitrary number, are listed and described in table 4. Their inherent contrasts (as positioned on the terrain model) were previously obtained, using a photopic luminosity criterion (Ref. 18).

Table 4. TARGETS USED IN SECOND DYNAMIC IMAGERY EXPERIMENT

| Target Number | Target | Size (Length x Width, ft.) | | Inherent Contrast |
|---|---|---|---|---|
| 1 | Convoy of 5 Missile Vans | 37 | 15 | 0.189 |
| 2 | 6 Ammo Bunkers | 55 | 30 | 0.414 |
| 4 | 11 Unit Train | 85 | 21 | 0.375 |
| 8 | Railroad Yard | 3,990 | 2,236 | 0.279 |
| 10 | 5 Combat Tanks | 22 | 12 | 0.200 |
| 11 | 5 Migs | 55 | 37 | 0.369 |
| 14 | Construction Yard | 1,000 | 875 | 0.550 |
| 16 | 6 POL Tanks | 75 Diameter | | 0.647 |
| 19 | Bridge | 386 | 25 | 0.122 |
| 21 | 6 POL Tanks | 75 Diameter | | 0.603 |
| 22 | 3 Large Buildings | 70 | 60 | 0.559 |
| 23 | 3 Boats | 90 | 25 | 0.234 |
| 25 | 4 Migs | 55 | 37 | 0.369 |
| 26 | Airport | 4,212 | 792 | 0.401 |
| 31 | 5 Small Buildings | 40 | 30 | 0.144 |
| 34 | Intersection of 2 Roads | 520 | 310 | (not definable) |
| 36 | SAM Site | 340 Diameter | | 0.414 |
| 38 | 1 Bridge and 3 Boats | 684 | 25 | 0.500 |
| 40 | 6 Ammo Bunkers | 66 | 32 | 0.662 |
| 42 | Construction Yard | 1,000 | 875 | 0.550 |
| 44 | 6 POL Tanks | 75 Diameter | | 0.286 |
| 45 | Construction Yard | 1,000 | 875 | 0.632 |
| 47 | SAM Site | 340 Diameter | | 0.414 |
| 49 | Harbor Complex | 2,170 | 1,396 | 0.167 |
| 51 | 4 Small Buildings | 40 | 50 | 0.324 |

PROCEDURE

Upon stating that he was familiar with each of the targets in some detail, the subject was placed in the experimental seat, which was adjusted for comfort,

with his eyes 40 inches from the 17-inch (diagonal) monitor, and with his head resting against a cushioned forehead bar. In the subject's lap was the target photo book, illuminated by a dim, but adequate, floodlamp in such a manner that there was no glare from this lamp into the subject's eyes as he viewed the monitor. He was easily able to look at both the photo book and at the monitor without appreciably moving his head. In one hand he held a response button which was used to signal when he recognized the prebriefed target. When the subject responded, indicating that he recognized the target, he also verbally indicated in which fourth of the display the target was located at the time he responded. For his convenience, the display vertical sides had tape markers placed to divide the display into four equal horizontal slices, so that the subject would respond "one" when the target was in the top fourth of the display, etc. Instructions to the subjects emphasized a "recognition" criterion. Subjects were permitted to cancel an erroneous response by responding again if necessary.

When the subject was ready, a static slide was presented on the monitor to indicate the general appearance of the terrain that he would be viewing, the scale factor involved, and the noise level of the video. He was then instructed to turn his photo book to the first target and begin searching for it when the terrain image on the monitor reappeared. The film was threaded into the projector, the counter was zeroed, the subject was cued to begin, and the trial was started. The experimenter informed the subject, via an intercom, when each target passed out of the field of view, so that he should begin searching for the next target in the photo book. In this manner, the total of 25 targets was presented to each subject in approximately 45 minutes.

Prior to running each subject, the video level of the system was checked, and the monitor was adjusted for tolerance to a gray scale input as indicated in section III. During each trial, the video level was monitored, although the automatic gain control of the CCU eliminated the need for any adjustments during the trial.

## OBSERVER PERFORMANCE RESULTS

Each subject's response to each target was scored as correct, incorrect, or no response. A correct response was one which was made at a frame number bounded

by 30 frames before the frame at which the target entered the designated fourth of the display and 30 frames after the target left the designated fourth of the display. Each target was on the total display for approximately 1345 frames, or 44.8 seconds. Thus, with as many as 2000 frames between successive targets, the likelihood of a subject guessing when a target was within a given fourth of the display was very small; therefore, no correction for guessing was made.

If a subject made more than one response to a given target, the last response was the one used for scoring purposes. All prior responses were ignored.

Responses were scored in terms of (1) proportion of targets correctly recognized, (2) proportion of targets incorrectly responded to, and (3) slant range to the target at the time of a response. Table 5 shows the values of these measures, averaged across subjects and targets, for each of the five noise levels.

Table 5. RESPONSE VALUES FOR SECOND DYNAMIC IMAGERY EXPERIMENT

| Noise Level, $\sigma_N$ | Proportion Correct | Proportion Incorrect | Mean Slant Range (feet) |
|---|---|---|---|
| 0 | .66 | .14 | 20,071 |
| .006 | .54 | .14 | 19,996 |
| .013 | .46 | .21 | 19,803 |
| .020 | .37 | .26 | 20,039 |
| .027 | .34 | .32 | 19,029 |

Correct vs. Incorrect Responses

Table 6 shows the results of an analysis of variance on the number of correct responses vs. the number of incorrect responses across the five noise levels. As illustrated in figure 37, the proportion of correct responses decreased with increases in noise, while the proportion of incorrect responses increased with increases in noise, as indicated by the statistically significant Noise x Correct vs. Incorrect interaction ($p < .001$). The difference between the total

proportion correct and the total proportion incorrect is also significant ($p < .001$). For convenience, figure 37 also shows the proportion of targets to which no response was made, although this is not included in the statistical analysis because the three proportions must necessarily sum to 1.00 and are therefore not independent.

Of little interest is the Sex x Noise interaction which, though statistically significant ($p < .001$), indicates a very slight crossover at the middle noise levels for the means of the two Sex levels.

Slant Range

Each response, correct or incorrect, was converted to the slant range to the target at the time of response. Then, for each subject the slant ranges of all correct responses were averaged to obtain one value, and the slant ranges of all incorrect responses were separately averaged to obtain another value. These single values were subjected to an analysis of variance, the results of which are summarized in table 7.

The mean slant range for all incorrect responses is larger than for all correct responses ($p < .001$). No other differences are significant, including the main Noise effect and all interactions involving Noise. The mean incorrect slant range is 23,027 feet, while the mean correct slant range is 19,788 feet, for a mean difference of 3,239 feet. The order of this difference is consistent at all noise levels, as illustrated in figure 38. Apparently, incorrect responses are typically made before the target is recognizable on the display - that is, while the target is still out of the field of view or before it is sufficiently large to recognize. In the latter event, another (non-target) object is responded to by the subject. When a corrective (second) response is made, the first response was ignored in the scoring. When, however, no second response was made, the first response slant range was used for incorrect slant range calculation. Thus, this difference between mean correct and mean incorrect slant range may be somewhat spurious.

Table 6. ANALYSIS OF VARIANCE OF NUMBER CORRECT VS. INCORRECT

| Source | SS | df | MS | F |
|---|---|---|---|---|
| Noise(N) | 62.39 | 4 | 15.60 | 5.53 * |
| Sex(S) | 0.09 | 1 | 0.09 | 0.03 |
| S x N | 68.24 | 4 | 17.06 | 6.05 ** |
| Subjects within S,N (Ss/S,N) | 126.90 | 45 | 2.82 | |
| Correct vs. Incorrect (C) | 1122.76 | 1 | 1122.76 | 90.25 ** |
| C x N | 590.86 | 4 | 147.71 | 11.87 ** |
| C x S | 10.05 | 1 | 10.05 | 0.81 |
| C x S x N | 58.34 | 4 | 14.58 | 1.17 |
| C x Ss/S,N | 559.83 | 45 | 12.44 | |
| | 2599.46 | 109 | | |

* $p < .01$

** $p < .001$

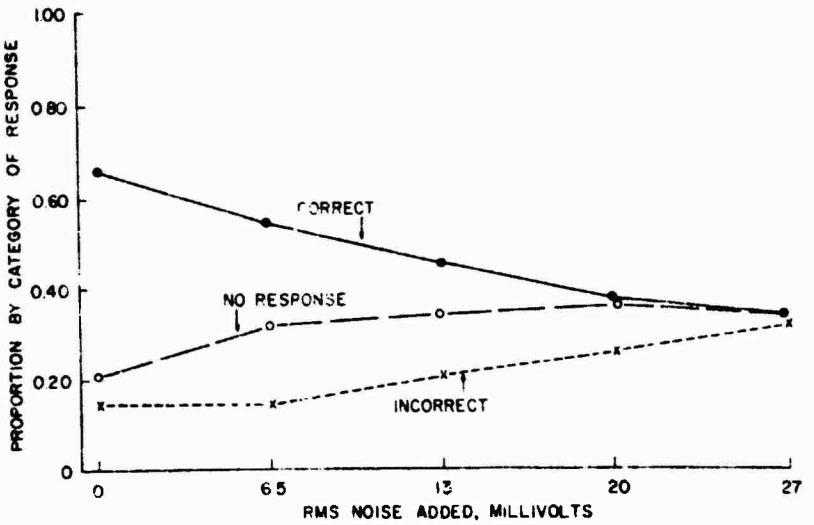



Figure 37. Proportion Correct Responses, Incorrect Responses, and No Responses as a Function of Noise Level.

Table 7. ANALYSIS OF VARIANCE SUMMARY, SLANT RANGE DATA, ONE SCORE PER SUBJECT

| Score | SS | df | MS | F |
|---|---|---|---|---|
| Noise Level(N) | 25,720,527 | 4 | 6,430,132 | 0.71 |
| Sex(S) | 216,277 | 1 | 216,277 | 0.02 |
| S x N | 85,426,717 | 4 | 21,356,679 | 2.34 |
| Subjects (Ss/S,N) | 410,136,764 | 45 | 9,114,150 | |
| Correct vs. Incorrect(C) | 288,603,006 | 1 | 288,603,006 | 33.54 ** |
| C x N | 83,125,373 | 4 | 20,781,343 | 2.41 |
| C x S | 1,691,406 | 1 | 1,691,406 | 0.20 |
| C x N x S | 68,104,657 | 4 | 17,026,164 | 1.98 |
| C x Ss/S,N | 378,547,100 | 44 * | | |
| | 1,341,571,827 | 108 * | | |

* -1 df for missing data

** $p < .001$

MEASUREMENT OF IMAGE QUALITY

The summary measures of image quality to be considered at this time are $N_e$, $SNR_D$, and MTFA, as discussed in section I. Each of these measures requires knowledge of the sine-wave response, R(N), of the system in the absence of noise. To obtain this measures, it was decided to use the square-wave response and transform analytically to the equivalent sine-wave response (Ref. 19), if necessary, for the following reasons. First, it is very difficult to obtain sine-wave targets in a variety of spatial frequencies and modulations in 35mm. transparency format of the type required to insert into the gate of the 35mm. film projector used in this experiment. Using any other format, and looking at the target through another optical system would not provide precise measurement of the sine-wave response of the equipment as used in this experiment. Second, of those sine-wave target generation techniques available, there seems to be some problem of actually producing a true sinusoidal modulation. Thus, the inherent

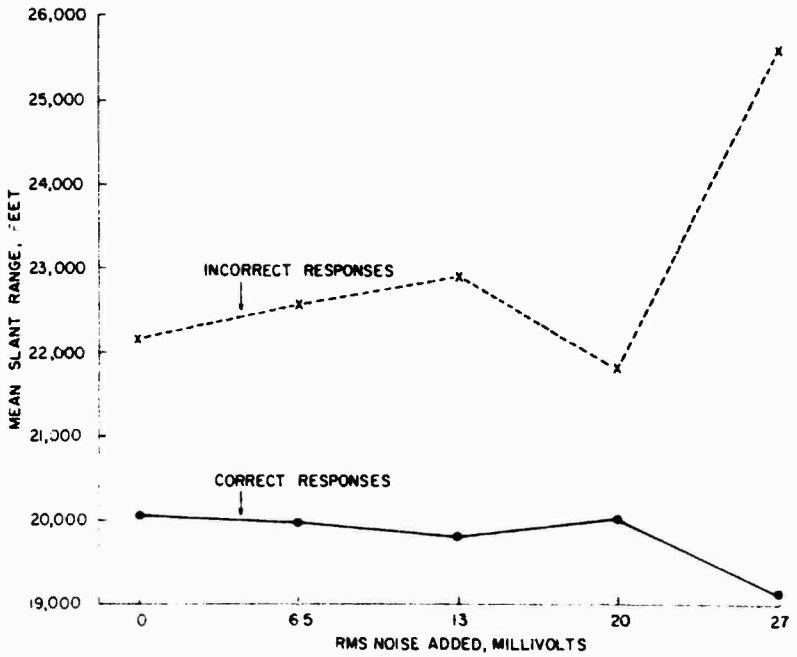



Figure 38. Effect of Noise Level on Slant Range For Correct and Incorrect Responses.

accuracy, repeatibility, and measurability of the square-wave target have several advantages. A repetitive strip of the Standard 1951 USAF tri-bar target was obtained on 35mm film and checked with the microdensitometer for true spatial frequency and modulation. A microdensitometer scanning spot of 20μ was used, and indicated that the square-wave modulation was 100% to a spatial frequency beyond the displayed monitor equivalent of 100 TV lines per inch.

The tri-bar target pattern was inserted into the projector film gate, centered so that the spatial frequency of interest was at the approximate center of the monitor with the target bars perpendicular to the TV raster, and measured at the monitor with a scanning microphotometer. The microphotometer has a scanning eyepiece with a movable slit of 25μ by 2500μ. The total scanning distance of the eyepiece is 10mm, which, with an objective of unity power, required repositioning of the eyepiece in order to scan all three black bars of the larger targets. The microphotometer output was recorded directly on an X-Y plotter, and the depth of modulation, in luminance units, was measured from the tracing. The square-wave response factor $R_{sq}(N)$ was then calculated for each target spatial frequency, with the resultant curve shown in figure 39. Also indicated in figure 39 is the

calculated sine-wave response, assuming linearity and based upon the approach indicated by Scott (Ref. 19), as well as the curve $[R(N)]^2$, from which $N_e$ can be calculated.

MTFA and System Performance

The MTFA concept requires that the system sine-wave response is known, and that detectability threshold curves, based upon this sine-wave response, are used (section I of this report). For reasons given previously, the detectability threshold curves were obtained for square-wave targets; thus, the modified concept $MTFA_{SQ}$, based upon both square-wave targets and system square-wave response, $R_{sq}(N)$, is used in these analyses. The $MTFA_{SQ}$ is simply the area bounded by the $R_{sq}(N)$ curve and the appropriate threshold curve, figure 39. Also, the threshold curves used here are not the same as those shown in section III of this report because a separate set of threshold curves was empirically obtained in a pilot experiment prior to the conduct of the research discussed in section III. These threshold curves are also shown in figure 39 for the 5 noise levels of this experiment. Figure 40 illustrates the relationship between $MTFA_{SQ}$ and the proportion of targets correctly recognized for the five noise levels. The product-moment correlation, based on these five points, is 0.965 ($p < .01$).

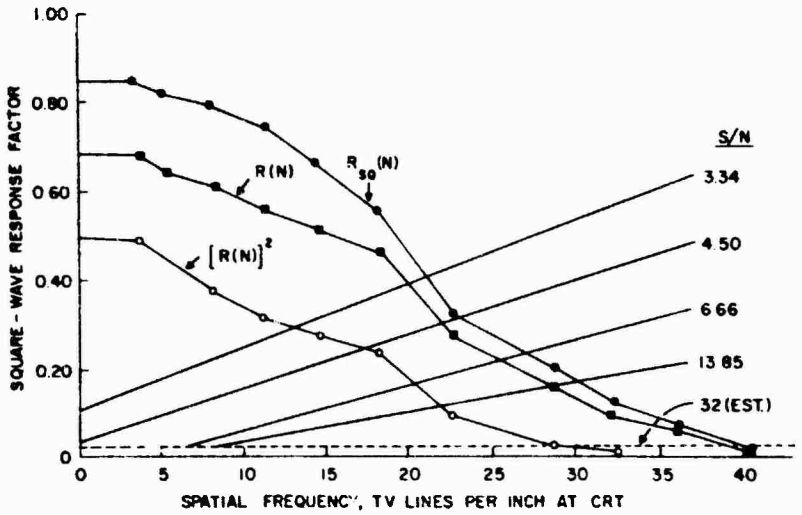



Figure 39. Square-Wave Response, Calculated Sine-Wave Response, and Threshold Curves of System.

The relationship between the proportion of incorrect responses and $MTFA_{SQ}$ is also shown in Figure 40, with its correlation of -0.973, which is also significant for df = 3 ($p < .01$).

The correlation of 0.765 between $MTFA_{SQ}$ and slant range for correct responses is shown in figure 41. It is not statistically significant, although the direction of the correlation is as expected.

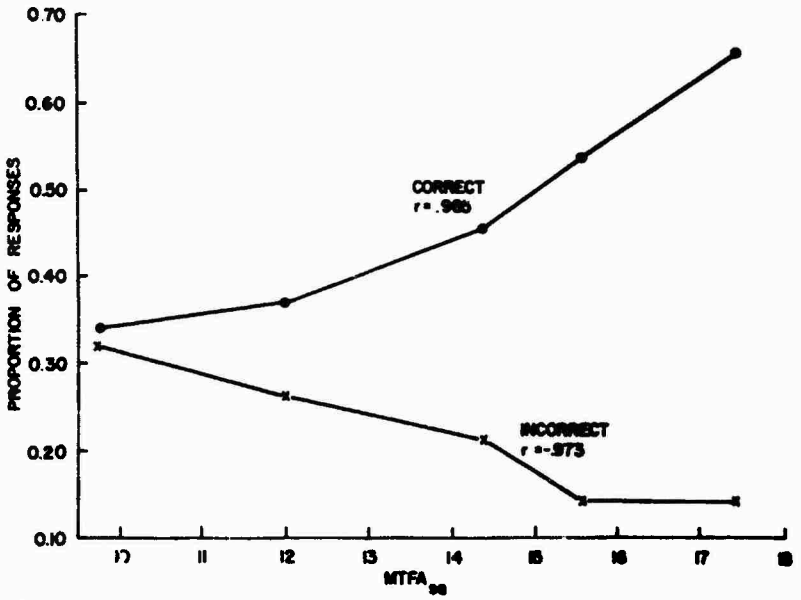



Figure 40. Prediction of Correct and Incorrect Responses by $MTFA_{SQ}$.

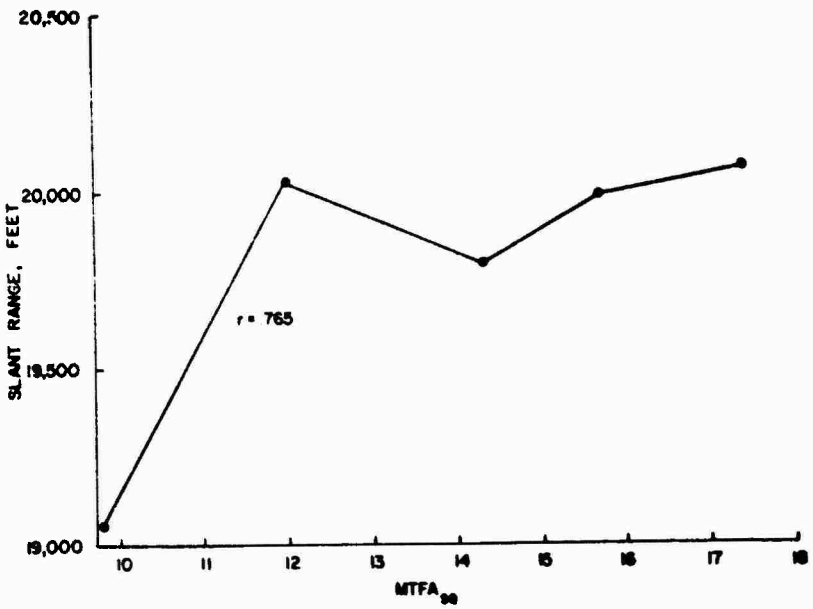



Figure 41. Prediction of Slant Range by $MTFA_{SQ}$.

## $SNR_D$ and System Performance

The evaluation of the $SNR_D$ concept for the present experiment requires a choice of formula for the calculation of the $SNR_D$, largely because the $SNR_D$ measure assumes that the area of the target is known (a in equation 3, section I). Averaging across the 25 targets, however, one can assume an unknown, but constant, average for a for any given noise level, and thereby calculate an average $SNR_D$ from:

$$SNR_D \sim \frac{\text{p-p signal}}{\text{rms noise}} \tag{10}$$

because all other terms in the $SNR_D$ formulae are constants for a given system. After performing these calculations, the values of (p-p signal)/rms noise are those previously given in table 3.

Correlations between this calculated value of signal/noise and the several measures of observer performance are 0.968 ($p < .05$) between S/N and percent correct recognition, -0.817 between S/N and percent incorrect recognition, and 0.514 between S/N and mean correct recognition slant range. The last two correlations are not statistically significant due to the small (i.e.,3) degrees of freedom. Figures 42 and 43 illustrate these correlations.

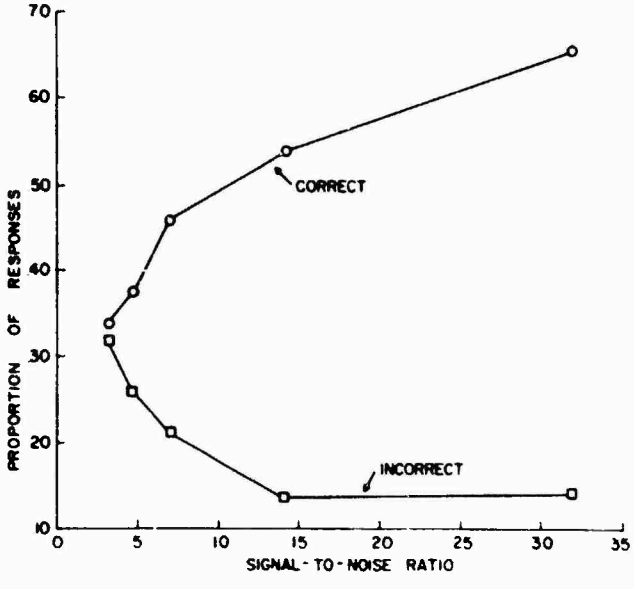



Figure 42. Prediction of S/N Correct and Incorrect Response by S/N.

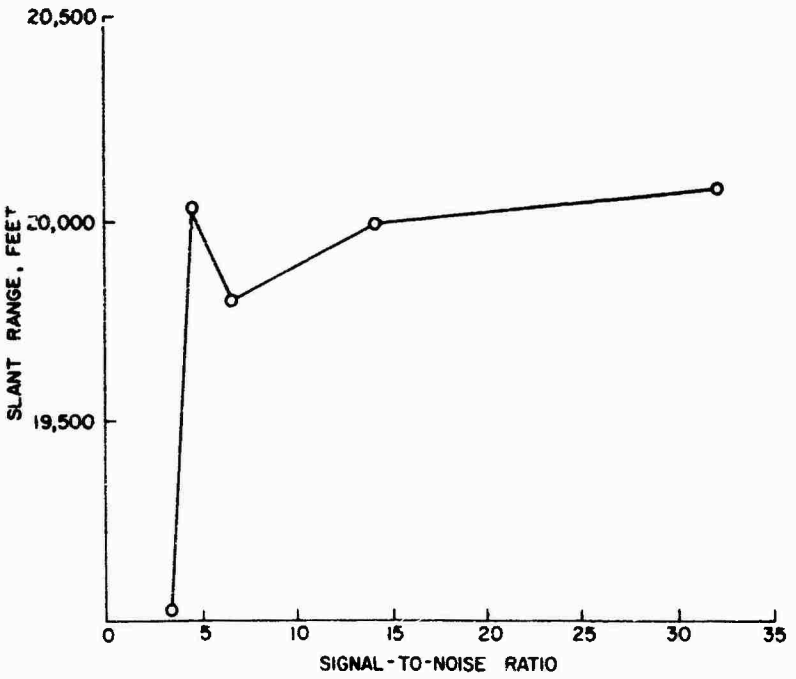



Figure 43. Prediction of Slant Range by S/N.

$N_e$ and System Performance

The use of the $N_e$ concept is not appropriate for comparing varying noise levels simply because the value of $N_e$ is dependent only upon the squared sine-wave response $[R(N)]^2$ of the system in the absence of noise, and takes no account of the varying noise levels in the system. The concept permits the calculation of the noise power (or voltage) transmitted by a given system amplifier through the particular system MTF, but does not take into account the observer's characteristics or needs for contrast, as do the $MTFA_{SQ}$ or $SNR_D$ approaches.

APPLICATION TO INDIVIDUAL TARGET PERFORMANCE PREDICTION

Both the $MTFA_{SQ}$ and $SNR_D$ metrics include a term pertaining to the contrast or modulation of the target object against its background. In addition, the $SNR_D$ formulae include a term (a/A) for the area of the target proportional to the total field of view, while the $MTFA_{SQ}$ ignores target size in favor of integrating

with equal weight over all spatial frequencies. The inclusion of such terms makes it possible to determine the ability of either metric to predict recognition performance on a target-by-target basis, rather than on a system comparison basis. A first estimate of such predictive ability will be presented in this section, while a more thorough examination of this problem will be included in section VI of this report.

## $MTFA_{SQ}$ and Target Prediction

The MTFA measure, as described in section I, includes an adjustment of the threshold detectability curve for the inherent target modulation. Such an adjustment assumes that the ordinate of any MTFA plot is the modulation transfer factor value, and that the threshold curve plotted is the required modulation of the target object for a threshold response. With the modification to the MTFA concept employed in this research to convert the MTFA sine-wave metric to $MTFA_{SQ}$, an additional modification has been employed; that is, the ordinate of the $MTFA_{SQ}$ plot has been changed to "square-wave modulation, CRT" or displayed square-wave modulation. This appears to be a more consistent label in that both the system response curve, $R_{sq}(N)$, and the threshold curves are now in common terms - the displayed modulation. Following this rationale, the adaptation of the MTFA approach to individual target prediction is to multiply the system $R_{sq}(N)$ curve by the inherent target/background modulation, thereby lowering the $R_{sq}(N)$ curve proportionately for targets of inherent modulation less than unity. Stated another way, the displayed target modulation is further reduced by the $R_{sq}(N)$ curve as the target becomes smaller.

This adjustment of the system response curve was made using part of the data to be subsequently described in section VI of this report, specifically the mean luminance of the target as compared with the mean luminance of the background on either side of the target to a point 25% of the target's width.

For reasons to be discussed in section VI, the inherent modulation measure, $M_o$, was obtainable for only 21 of the 25 targets, excluding targets numbered 22, 23, 31, and 51 in table 4. For the remaining 21 targets, the $R_{sq}(N)$ curve was multiplied by the target's $M_o$ value, and an $MTFA_{SQ}$ was calculated for each target for each of the five noise levels, or a total of 105 $MTFA_{SQ}$ values. The

TABLE 3. SUMMARY DATA FOR $MTFA_{SQ}$, BY TARGET, AND OBSERVER TARGET RECOGNITION PERFORMANCE.

| | | $MTFA_{SQ}$ by Noise Level | | | | | Percent correct by Noise Level | | | | | Mean Slant Range by Noise Level | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Target Number | $M_0$ | 0 | .006 | .013 | .020 | .027 | 0 | .006 | .013 | .020 | .027 | 0 | .006 | .013 | .020 | .027 |
| 1 | .167 | 2.15 | 1.48 | 1.17 | 0.40 | 0.06 | 63.6 | 30.0 | 27.3 | 18.2 | 0 | 17996 | 17446 | 19496 | 18883 | - |
| 2 | .520 | 8.10 | 6.57 | 5.86 | 4.26 | 2.86 | 81.8 | 27.3 | 0 | 0 | 0 | 18070 | 17607 | - | - | - |
| 4 | .091 | 0.89 | 0.58 | 0.44 | 0.08 | 0 | 9.1 | 10.0 | 27.3 | 0 | 27.3 | 17435 | 17432 | 21063 | - | 19596 |
| 8 | .536 | 8.10 | 6.57 | 5.86 | 4.26 | 2.86 | 81.8 | 90.9 | 54.5 | 72.7 | 54.5 | 21299 | 19726 | 18998 | 18302 | 17733 |
| 10 | .167 | 2.15 | 1.48 | 1.17 | .40 | .06 | 9.1 | 9.1 | 0 | 9.1 | 9.1 | 18025 | 20211 | - | 17625 | 22721 |
| 11 | .258 | 3.63 | 2.76 | 2.29 | 1.21 | 0.44 | 9.1 | 0.0 | 0 | 0 | 9.1 | 17548 | - | - | - | 17754 |
| 14 | .203 | 2.77 | 1.99 | 1.66 | 0.69 | 0.16 | 100. | 72.1 | 81.8 | 72.7 | 54.5 | 21280 | 21365 | 21099 | 21894 | 20237 |
| 16 | .632 | 8.67 | 8.07 | 7.24 | 4.93 | 3.39 | 100. | 100. | 81.8 | 63.6 | 72.7 | 23799 | 20743 | 20584 | 20294 | 18414 |
| 19 | .530 | 8.10 | 6.57 | 5.86 | 4.26 | 2.86 | 63.6 | 54.5 | 36.4 | 9.1 | 18.2 | 19851 | 18044 | 17729 | 17380 | 17475 |
| 21 | .623 | 8.67 | 8.07 | 7.24 | 4.93 | 3.39 | 81.8 | 100. | 72.7 | 72.7 | 72.7 | 21264 | 21540 | 22052 | 20683 | 18545 |
| 25 | .250 | 3.63 | 2.76 | 2.29 | 1.21 | 0.44 | 70.0 | 20.0 | 0 | 0 | 0 | 17762 | 17689 | - | - | - |
| 26 | .550 | 10.35 | 7.60 | 6.84 | 5.10 | 3.52 | 100. | 90.9 | 100. | 100. | 90.9 | 22937 | 22804 | 22730 | 23154 | 22276 |
| 34 | .371 | 5.40 | 4.22 | 3.72 | 2.22 | 1.28 | 81.8 | 81.8 | 90.9 | 63.6 | 54.5 | 18624 | 19635 | 18707 | 19355 | 17464 |
| 36 | .350 | 5.40 | 4.22 | 3.72 | 2.22 | 1.28 | 90.9 | 70.0 | 70.0 | 27.3 | 18.2 | 19647 | 18517 | 19136 | 19776 | 17385 |
| 38 | .286 | 4.27 | 3.27 | 2.82 | 1.49 | 0.67 | 100. | 63.6 | 63.6 | 27.3 | 18.2 | 20099 | 18812 | 18022 | 18448 | 18978 |
| 40 | .467 | 7.52 | 6.09 | 5.39 | 3.78 | 2.35 | 60.0 | 70.0 | 50.0 | 9.1 | 0 | 18982 | 18543 | 17941 | 18474 | - |
| 42 | .200 | 2.77 | 1.99 | 1.66 | 0.69 | 0.16 | 90.9 | 81.8 | 90.9 | 72.7 | 90.9 | 21593 | 20675 | 20035 | 20476 | 19169 |
| 44 | .570 | 10.35 | 7.60 | 6.84 | 5.10 | 3.52 | 90.9 | 70.0 | 70.0 | 81.8 | 63.6 | 20980 | 20572 | 19081 | 18560 | 18146 |
| 45 | .429 | 6.94 | 5.62 | 4.88 | 3.39 | 1.91 | 90.9 | 90.9 | 72.7 | 63.6 | 72.7 | 21169 | 20804 | 19430 | 19878 | 18571 |
| 47 | .269 | 3.63 | 2.76 | 2.29 | 1.21 | 0.44 | 60.0 | 30.0 | 0 | 0 | 0 | 17791 | 17480 | - | - | - |
| 49 | .571 | 10.35 | 7.60 | 6.84 | 5.10 | 3.52 | 100. | 72.7 | 63.6 | 100. | 63.6 | 20772 | 21165 | 20382 | 20263 | 19316 |

resultant $MTFA_{SQ}$ values, along with the observer performance data with which the $MTFA_{SQ}$ values are correlated, are presented in table 8.

The obtained correlations between the various performance measures and the by-target $MTFA_{SQ}$ are given in table 9. The correlations between $MTFA_{SQ}$ and the precent correct recognition measures vary from .411 to .651 for the five noise levels, indicating that the $MTFA_{SQ}$ measure is a reasonable predictor of the likelihood of recognizing a particular target with a given system noise level, but that other parameters are also important. Section VI will explore other target and background parameters.

Similarly, the correlations between $MTFA_{SQ}$ and slant range vary from -.185 to .597 over the five noise levels. Obviously, the $MTFA_{SQ}$ metric is not a very consistent predictor of slant range at the time of recognition.

$SNR_D$ and Target Prediction

Equation (2) in section I related $SNR_D$ to the variables a (target area on photocathode), C (target contrast), and $i_{max}$ (the maximum photocurrent), plus a few other variables which were held constant in this experiment. To assess the ability of $SNR_D$ to predict target recognition performance on a target-by-target basis, the following formula was used to establish $SNR_D'$, which is proportional to $SNR_D$.

$$SNR'_D = (a)^{1/2} \left( \frac{C - (\text{Signal, p-p})}{\text{Noise, rms}} \right)$$

$$= (a)^{1/2} \left( \frac{0.090\ C}{\text{Noise, rms volts}} \right) \qquad (11)$$

where a = target area
C = target/background contrast, defined by

$$\frac{\text{Luminance of target - Luminance of background}}{\text{Maximum luminance, target or background}}$$

Table 9. PRODUCT MOMENT CORRELATION OF $MTFA_{SQ}$ BY TARGET WITH OBSERVER PERFORMANCE

| Noise Level | $r$($MTFA_{SQ}$, Percent Correct) | $r$($MTFA_{SQ}$, Slant Range) |
|---|---|---|
| 0 | 0.556 ** | 0.597 ** |
| .006 | 0.651 ** | 0.451 * |
| .013 | 0.411 | 0.096 |
| .020 | 0.527 * | 0.148 |
| .027 | 0.452 | -0.185 |

*$p < .05$
**$p < .01$

The values of the calculated $SNR_D'$ are given in table 10, while table 11 shows the correlations between $SNR_D'$ and both percent correct recognition and slant range. These correlations range from 0.380 to 0.663 for prediction of the percent targets correctly recognized, and from 0.121 to 0.520 for the slant range at the time of correct recognition.

## DISCUSSION

The results clearly indicate that either $MTFA_{SQ}$ or $SNR_D$ is a reasonably good predictor of overall system performance, as measured by the proportion of correct responses, the proportion of incorrect responses, or the slant range at the time of a correct response. Differences between $MTFA_{SQ}$ and the $SNR_D$ predictors are generally small, although the numerical correlations are higher for the $MTFA_{SQ}$ measure. The similarity of the magnitude of these correlations is not very surprising inasmuch as the two measures are very similar in concept and, under certain conditions, equivalent (ref. 2), as will be discussed later.

$N_e$ is simply not an appropriate metric to use for comparisons of system performance under varying noise levels, as in this experiment.

Table 10. SUMMARY DATA FOR PREDICITION OF TARGET-BY-TARGET RECOGNITION PERFORMANCE FROM $SNR_D'$

| Target Number | Target Area Ground Units, ft.$^2$ | C | 0 | $SNR_D'$ By Noise Level .006 | .013 | .020 | .027 |
|---|---|---|---|---|---|---|---|
| 1 | 2,775 | 0.71 | 1196.85 | 258.82 | 124.55 | 84.15 | 62.46 |
| 2 | 26,889 | 0.62 | 3253.34 | 703.53 | 338.53 | 228.75 | 169.78 |
| 4 | 14,175 | 0.73 | 2781.21 | 601.44 | 289.42 | 195.55 | 145.14 |
| 8 | 8,921,640 | 0.62 | 59260.32 | 12815.04 | 6166.78 | 4166.74 | 3092.65 |
| 10 | [illegible],225 | 0.71 | 3067.20 | 663.28 | 319.18 | 215.66 | 160.07 |
| 11 | 46,035 | 0.69 | 4737.43 | 1024.47 | 492.99 | 333.10 | 247.23 |
| 14 | 875,000 | 0.70 | 20953.28 | 4531.15 | 2180.45 | 1473.28 | 1093.50 |
| 16 | 84,375 | 0.59 | 5484.14 | 1185.95 | 570.69 | 385.60 | 286.20 |
| 19 | 9,650 | 0.62 | 1948.97 | 421.46 | 202.81 | 137.04 | 101.71 |
| 21 | 84,375 | 0.59 | 5484.16 | 1185.95 | 570.69 | 385.60 | 286.20 |
| 25 | 11,884 | 0.61 | 2128.00 | 460.18 | 221.45 | 149.63 | 111.06 |
| 26 | 3,335,904 | 0.61 | 35652.16 | 7709.78 | 3710.05 | 2506.79 | 1860.60 |
| 34 | 161,200 | 0.66 | 8479.68 | 1833.73 | 882.42 | 596.23 | 442.53 |
| 36 | 90,792 | 0.66 | 6363.84 | 1376.18 | 662.24 | 447.46 | 332.11 |
| 38 | 17,100 | 0.68 | 2845.44 | 615.33 | 296.10 | 200.07 | 148.50 |
| 40 | 44,280 | 0.63 | 4242.24 | 917.38 | 441.46 | 298.28 | 221.39 |
| 42 | 875,000 | 0.70 | 20953.28 | 4531.15 | 2180.45 | 1473.28 | 1093.50 |
| 44 | 84,375 | 0.61 | 5670.08 | 1226.15 | 590.04 | 398.68 | 295.91 |
| 45 | 875,000 | 0.64 | 19157.12 | 4142.73 | 1993.54 | 1346.99 | 999.76 |
| 47 | 90,792 | 0.68 | 6556.80 | 1417.91 | 682.32 | 461.03 | 342.18 |
| 49 | 3,029,320 | 0.61 | 33974.40 | 7346.96 | 3535.46 | 2388.83 | 1773.04 |

The prediction of target recognition performance on a target-by-target basis is a totally different matter, however. As tables 9 and 11 show, both $MTFA_{SQ}$ and $SNR_D'$ predict individual target performance to only a small extent (correlations between -0.185 and 0.663, with a mean correlation of 0.400 for an average prediction of 16 percent of the target variance). Thus, the prediction of individual target recognition performance must take into account much more than the area of the target, the target's contrast with its background, and the characteristics of the imaging system. This result is not very surprising,

Table 11. CORRELATION OF $SNR_D'$, BY TARGET, WITH OBSERVER PERFORMANCE

| Noise Level | $r(SNR_D', \text{Percent Correct})$ | $r(SNR_D', \text{SLANT RANGE})$ |
|---|---|---|
| 0 | 0.380 | 0.520 * |
| .006 | 0.494 * | 0.520 * |
| .013 | 0.413 | 0.257 |
| .020 | 0.663 ** | 0.345 |
| .027 | 0.577 ** | 0.121 |

*$p < .05$
**$p < .01$

since several experiments in the past have investigated target complexity (refs. 21, 23, 24) and found that many additional parameters are involved. Another effort in this line of research will be discussed in section IV of this report.

As in the case of overall system comparisons involving noise, the $N_e$ metric is not usable to predict differences among targets, and therefore is not recommended as an overall useful measure of image quality for video systems.

## SECTION V

## A STATIC IMAGERY EXPERIMENT

The contents of this section, although not measuring air-to-ground target recognition, do relate to the problem of video system image quality and further demonstrate the utility of the $MTFA_{SQ}$ metric in the prediction of observer performance from a video display. This research was conducted as an "add-on" to the threshold research of section III, using the same subjects during unscheduled time of the equipment.

### INTRODUCTION

Experimental consideration was given to the problem of defining the requirements of an imaging system for police surveillance applications, where the electro-optical system is often used under various scene irradiance conditions, with various optical components, and to view various types of persons, vehicles, etc. Specifically, the ability of persons to identify static images of human faces was determined under a combination of noise levels and TV system configurations.

### EXPERIMENTAL DESIGN

To investigate the relationship between $MTFA_{SQ}$ and observer performance in facial recognition, a total of 15 different $MTFA_{SQ}$ values was generated by combining three television system $R_{sq}(N)$ curves with five signal-to-noise levels. Five threshold curves from section III were selected for each of the $R_{sq}(N)$ conditions, as shown in table 12 and figures 44-46. The five values for each condition were selected to provide the approximate same subjective noise levels for each $R_{sq}(N)$ condition, respectively. Integration of the areas bounded by the combinations of these three $R_{sq}(N)$ curves and the five detectability threshold functions for each of the $R_{sq}(N)$ curves was performed to obtain the $MTFA_{SQ}$ values indicated in table 12. Each of 5 observers was given a randomly selected subset of 7 different faces for each of the 15 $R_{sq}(N)$ signal-to-noise conditions. The order of presentation of each face within the seven-face subset was randomized. The only restrictions imposed upon the assignment of

faces to the subsets were that each of the 35 faces had to appear one time per subject per $R_{sq}(N)$ level, and that no subject experienced the same face under the same relative signal-to-noise conditions. Therefore, summing across the 5 subjects and the 35 trials per subject per $R_{sq}(N)$, each face was presented exactly 5 times, once per subject.

## EXPERIMENTAL EQUIPMENT

The equipment was essentially the same as that described in section II and as used in section III of this report. Instead of placing the tribar photographs in the viewing box, figure 10, the facial photographs were inserted. Each photograph was a "head and shoulders" picture, processed to the same mean gray level and gamma. Appropriate changes were made in the television camera and camera control unit to achieve the three $R_{sq}(N)$ conditions. Specifically, the $R_{sq}(N)$ curves were produced by combinations of line rate per frame and video passband, as follows: 8/525, 16/945, and 32/1225, where the first number is the video passband (-3dB) in megahertz, and the second is the line rate per frame.

The $R_{sq}(N)$ curves illustrated in figures 44-46 were obtained photometrically using tri-bar unputs of high modulation as described in section III. Prior to each experimental session, and every 30 minutes thereafter, the television system was recalibrated to assure constant gray-scale rendition and video gain. Details of this procedure and of the apparatus were presented in section III. Monitor mean luminance was fixed at 3 ft-Lamberts.

Table 12. EXPERIMENTAL DESIGN

| Bandwidth/ line rate | Noise rms, mV. | Subjects | | | | | $MTFA_{SQ}$ |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | |
| 8/525 | 0 | subset 1 | subset 2 | subset 3 | subset 4 | subset 5 | 30.46 |
| | 37 | subset 2 | subset 3 | subset 4 | subset 5 | subset 1 | 18.46 |
| | 50 | subset 3 | subset 4 | subset 5 | subset 1 | subset 2 | 13.22 |
| | 62 | subset 4 | subset 5 | subset 1 | subset 2 | subset 3 | 8.72 |
| | 75 | subset 5 | subset 1 | subset 2 | subset 3 | subset 4 | 4.48 |
| 16/945 | 0 | subset 6 | subset 7 | subset 8 | subset 9 | subset 10 | 21.09 |
| | 28 | subset 7 | subset 8 | subset 9 | subset 10 | subset 6 | 15.23 |
| | 42 | subset 8 | subset 9 | subset 10 | subset 6 | subset 7 | 11.47 |
| | 56 | subset 9 | subset 10 | subset 6 | subset 7 | subset 8 | 8.33 |
| | 70 | subset 10 | subset 6 | subset 7 | subset 8 | subset 9 | 5.64 |
| 32/1225 | 0 | subset 11 | subset 12 | subset 13 | subset 14 | subset 15 | 14.65 |
| | 40 | subset 12 | subset 13 | subset 14 | subset 15 | subset 11 | 7.37 |
| | 50 | subset 13 | subset 14 | subset 15 | subset 11 | subset 12 | 5.33 |
| | 60 | subset 14 | subset 15 | subset 11 | subset 12 | subset 13 | 3.67 |
| | 70 | subset 15 | subset 11 | subset 12 | subset 13 | subset 14 | 2.36 |

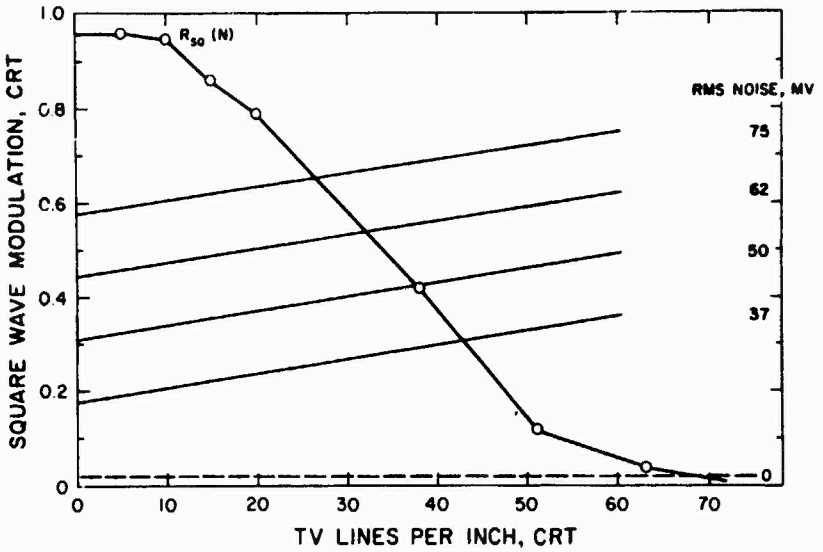



Figure 44. Square-wave Response and Threshold Curves for the 525-line, 8-megahertz System.

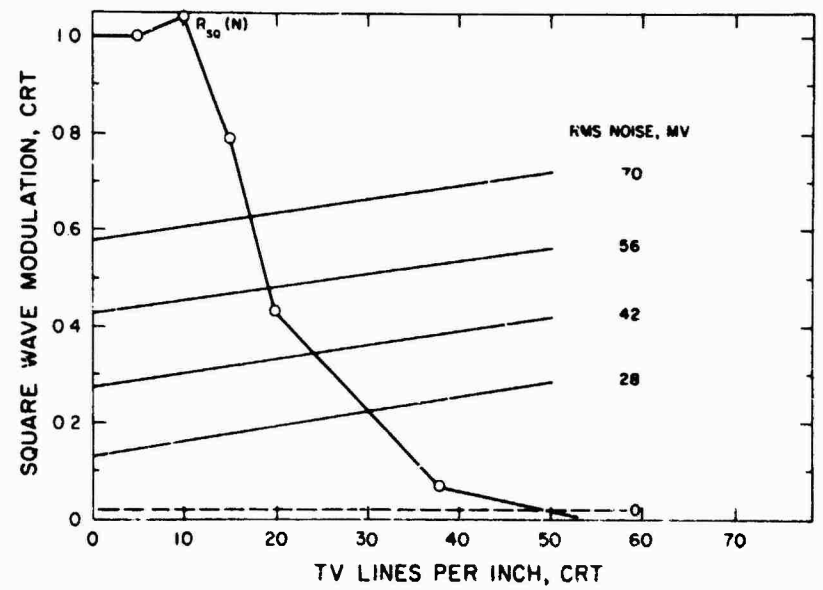



Figure 45. Square-wave Response and Threshold Curves for the 945-line, 16-megahertz System.

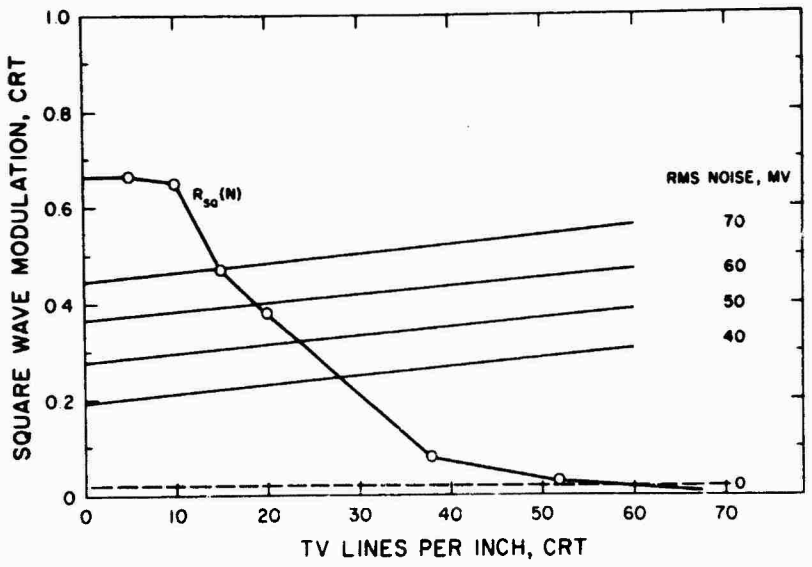



Figure 46. Square-wave Response and Threshold Curves for the 1225-line, 32-megahertz System.

Located approximately 24 inches to the left of the subject was a 31 x 48-inch wall-hung board containing, in random order, photographs of the 35 numbered faces to be viewed on the monitor. Each of these photographs was 4 x 5 inches in size, and was taken when the person posing for the photograph was wearing clothing different from that shown in the TV displayed photograph. In this manner, clothing cues were not present to assist the subject in recognizing individual faces. The wall-hung board was illuminated to a comfortable level with a small flood lamp.

Of the 35 faces, 2 were of females, and 1 of the 33 males was oriental, the rest caucasian. The sex difference did not appear to affect the subjects' responses, perhaps due to the similarity in hair styles among males and females in the photographs. All 35 photographs were of persons between 19 and 36 years of age.

## PROCEDURE

The subject was seated in an adjustable chair with his eyes approximately 40 inches from the vertically oriented 17-inch (diagonal) monitor. No room

lights were on. A single flood lamp illuminated the 35-face board to the subject's left. There was no reflection of this flood from either the individual photographs on the wall-hung board or the monitor.

A Standard Electric timer was started as each stimulus photograph was inserted into the holder in front of the television camera. When the subject recognized the stimulus photograph, he depressed a button that stopped the timer and he simultaneously stated the number of the face on the wall-hung gallery to his left. The correctness of his response was recorded along with the total response time to the nearest 0.01 second. Each subject was given all 35 photographs at a single $R_{sq}(N)$ level during one experimental session, typically lasting about 20 minutes. The 3 sessions per subject were spaced about 1 week apart.

## RESULTS

The data were scored in terms of two dependent variables, percent correct recognition and response time. The response time was read directly from the timer, while the percent correct recognition score was calculated as the number of correct recognition responses divided by the number of photographs presented. Inasmuch as each subject was forced to respond to each photograph, the denominator of the percent correct measure is 35 for each $R_{sq}(N)$/signal-to-noise combination.

Figure 47 shows the relation between the mean percent correct recognition for each of the 15 $R_{sq}(N)$ signal-to-noise conditions and $MTFA_{SQ}$, while figure 48 illustrates the relation between mean response time and $MTFA_{SQ}$. The linear correlation coefficients, 0.69 and -0.67, respectively, are both statistically significant, $p < 0.01$. However, it is apparent that a nonlinear relationship better describes the data.

A reasonably good fit to the percent correct recognition measure is shown in figure 49, where the $MTFA_{SQ}$ metric was transformed into $\log_{10}MTFA_{SQ}$. Using this transformation, the correlation between percent correct recognition and $\log_{10}MTFA_{SQ}$ is 0.87, which is significant at $p < 0.001$.

The response time data are better fit with a log-log transformation, as illustrated in figure 50, which results in a correlation of -0.92, $p < 0.001$.

While it is probably possible to fit more complex functions to these data to produce a better least-squares fit, the use of $\log_{10}$ transformations is reasonable in that the visual system, like many sensors, behaves largely in proportion to the logarithm of the energy impinging upon it. Further, previous research using the MTFA metric has found that a log or log-log transformation produces a good fit to this type of performance data (ref. 9). With the transformations indicated above, the least-squares best-fit equations become:

$$P_c = 0.4146 \log_{10} MTFA_{SQ} + 0.4688, \text{ and} \tag{12}$$

$$\log_{10} RT = 1.6658 - 0.7084 \log_{10} MTFA_{SQ} \tag{13}$$

$$\text{or } RT = 80.233\ MTFA_{SQ}^{-0.7084} \tag{13a}$$

where $P_c$ is the percent correct recognition, and

RT is mean response time

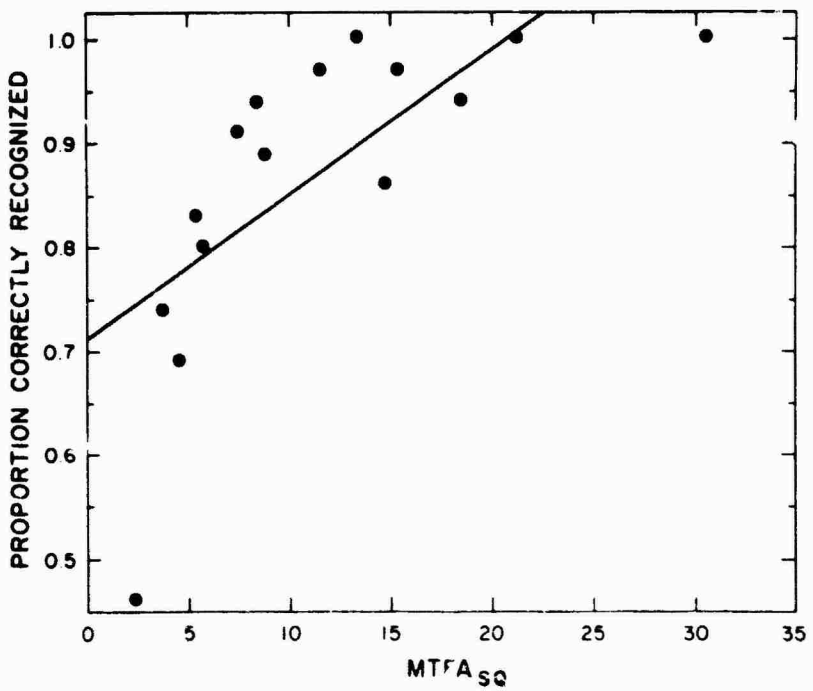



Figure 47. Proportion of Faces Correctly Recognized.

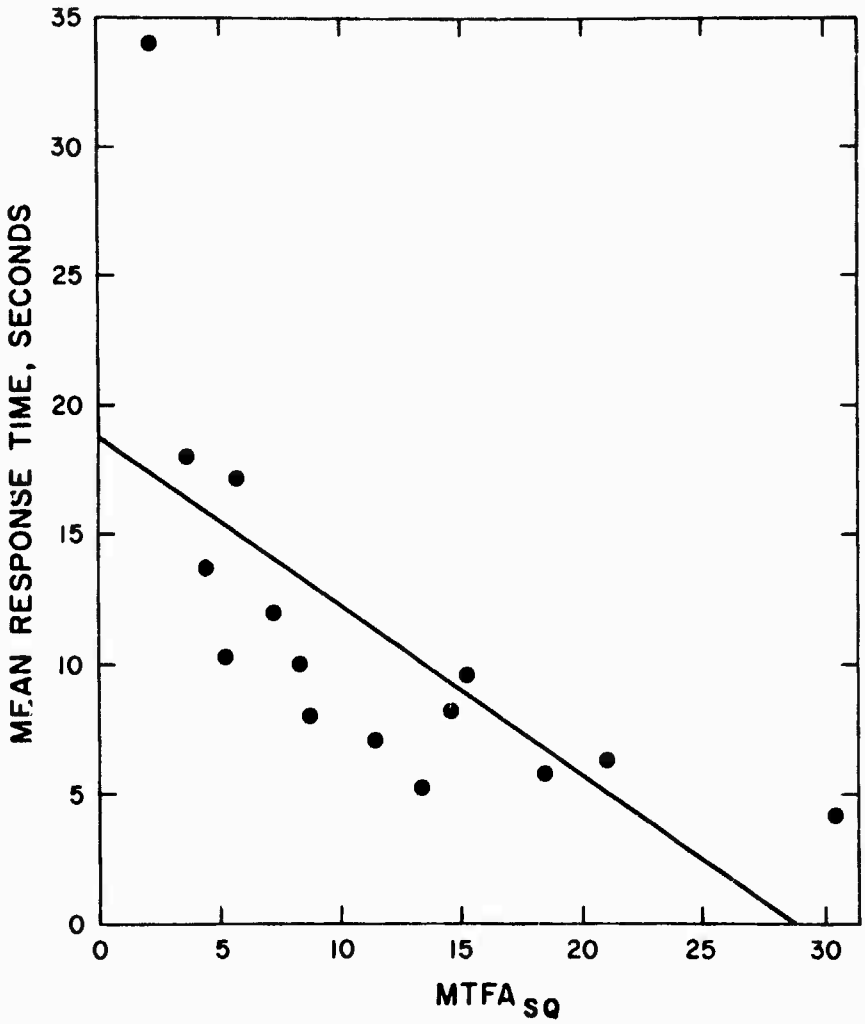



Figure 48. Mean Response Time for Facial Recognition.

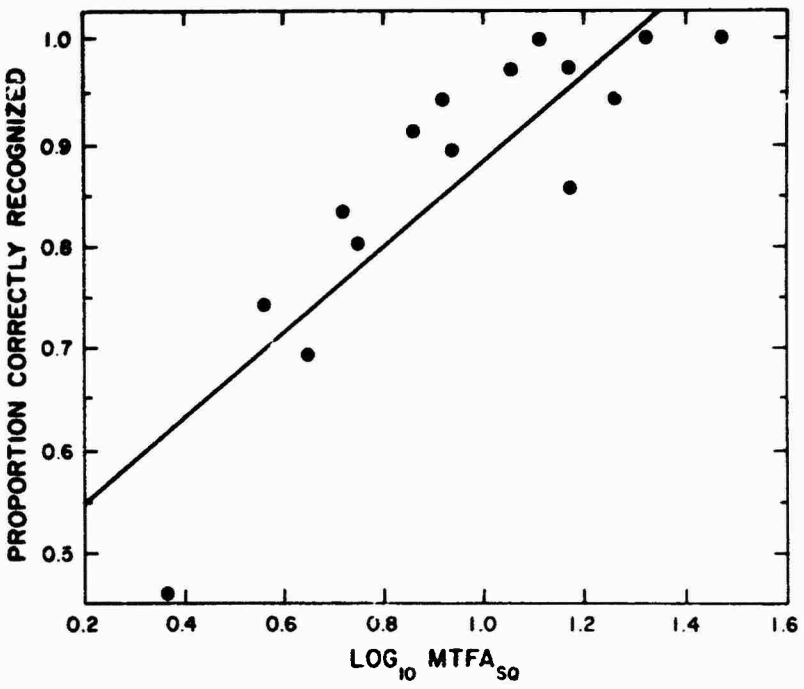



Figure 49. Proportion of Faces Correctly Recognized.

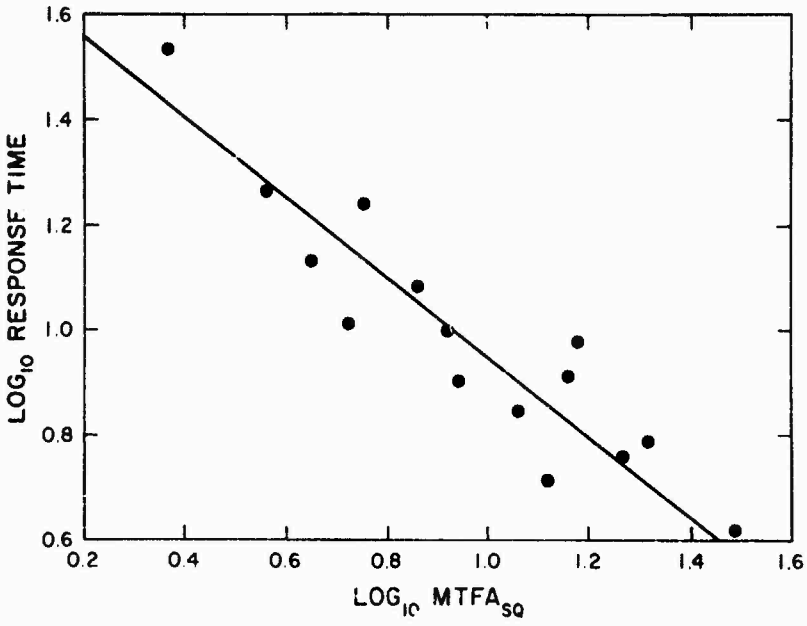



Figure 50. $Log_{10}$ Response Time for Facial Recognition.

## DISCUSSION

These results clearly indicate that the $MTFA_{SQ}$ metric of image quality is a strong predictor of the ability of persons to recognize static individual faces on a television display. This general result is in good agreement with data from section IV. That is, the $MTFA_{SQ}$ value appears to predict observer performance from both static and dynamic imagery.

With these particular data, a log or log-log transformation produces a better linear fit, although the linearity of the relationship is unimportant for estimating the magnitude of performance prediction. The nonlinearity of the best-fit expression in these data is probably due to the large facial images presented to the subjects (relative low spatial frequencies) and the simplicity of the task. That is, as the $MTFA_{SQ}$ value becomes even moderately large, facial recognition performance reaches a ceiling or asymptotic value which cannot be exceeded. Either the percent correct recognition approaches 100 or the mean response time approaches the minimum required for the subject to look from the television monitor to the gallery of faces to his left, find the single face of interest, and depress the response button. This response time lower limit seems to be about 5 seconds. When greater time was taken, it appeared to be because the subject both (1) studied the television image for a longer period of time, and (2) fixated alternately on the monitor and the face gallery several times, apparently in an effort to compare specific features of the face being presented on the television monitor with those of the face gallery. It is hypothesized that a useful index of image quality might be either the redundancy of eye fixation locations on the display or the mean time per fixation. Data to evaluate this notion are presently being taken in the second phase of this program.

## SECTION VI

## TARGET-BY-TARGET PREDICTION OF PERFORMANCE

Section IV results showed that the $MTFA_{SQ}$ and the averaged $SNR_D$ were excellent predictors of dynamic target acquisition performance across five alternate "systems" but that the same metrics were relatively poor predictors of the same performance measures on an individual target basis. It was pointed out that the $MTFA_{SQ}$ metric included information about the target/background contrast, and that the $SNR_D$ included information about both the target/background contrast and the size of the target. However, as several studies have shown in the past, these two target parameters do not adequately describe the "recognizability" of a target. Thus, it remains to define a suitable image quality metric which can be used to predict individual target recognition performance rather than overall system-average target recognition performance, simply because the operational commander has one prediction problem (the former) while the system designer has another problem (the latter). Based in part upon the recent work by Zaitzeff (ref. 21), this section describes a preliminary effort to predict individual target recognition performance on a video display directly from geometric and photometric knowledge of the target as might be available, for example, from a reconnaissance photograph.

The targets used were the same as those employed in the dynamic imagery experiment reported in section IV. Microdensitometric scans were made across each of the targets in the 35mm film frame. Various parameters of the microdensitometric scan trace were measured, along with the size of the target, to produce a total of 35 predictor variables. Stepwise linear multiple regression analysis was then used to predict each of four observer performance measures. The results indicate that perfect prediction is obtained with a maximum of 19 of the predictors for any one performance (criterion) variable, and that several of the 19 variables needed are common to prediction of the four performance measures. Also, a reduced set of 8 predictors predicts a major proportion of the variance in all four criterion variables.

## MICRODENSITOMETRY MEASURES

A Gamma Scientific microdensitometer was used to scan the 35mm. positive transparency frame of each of the targets listed in table 4, with the specific frame chosen to represent a slant range of 24,450 feet. A single horizontal scan was made across the entire frame at a point which was subjectively estimated to contain most of the distinguishing features of each target, for example, the road, control building, and missile launcher of a SAM site. The scanning aperture, at the transparency, was 60μ, which is considerably larger than the grain size of the print film. The output of the microdensitometer, in transmission units, was converted to equivalent ft.-Lamberts at the TV monitor from a previously obtained relationship using the 11-step gray scale as described in section III. This transmission output was recorded directly as Y on an X-Y plotter, with the X drive being the scanning platform output of the microdensitometer, or distance across the film frame.

During a second pass along the same line on the 35mm. frame, the transmission output of the microdensitometer was integrated, and the integral was recorded as Y on the X-Y plotter, with the X axis still representing horizontal position across the frame. Manual notation was made on the X-Y plot of the location of the target, key target elements, and various background elements so that both the transmission and integrated transmission traces could be easily referenced to scene content.

## PHOTOMETRIC MEASURES

Previous research (e.g., refs. 21, 22) employed both physical measures of the target and its background, and subjectively scaled measures obtained from a group of observers. While the merit of this approach should not be questioned in terms of its containing those variables of importance, it seems more desirable to include only objective measures which can be readily obtained under plausible field conditions, and which might be derived more or less automatically from a preprogrammed scanning apparatus, thus eliminating any subjective error in such measurement. For these reasons, this research restricted the predictor variables to only physical measurements derivable from either the geometry of

the target or from the two microdensitometric traces per target (both of which could be obtained from a single pass if the linear and integrated outputs were taken in parallel fashion).

Figure 51 illustrates a theoretical tracing across one film frame. The target, total background, and a section of the background equal to 25% of the target's width on either side of the target are indicated. Also shown is the integrated transmission curve which might be obtained for such a target. Figure 52 shows an actual tracing across a target.

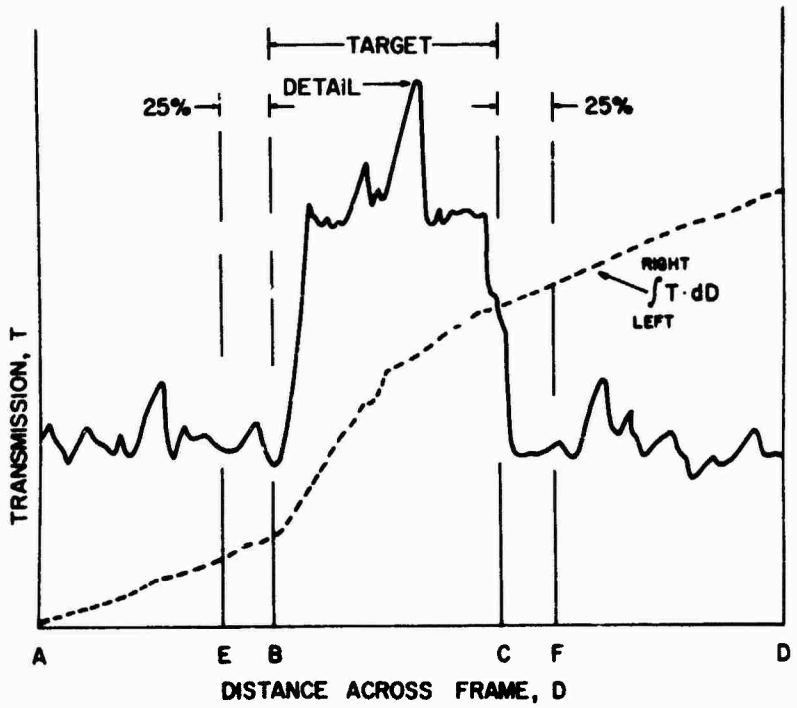
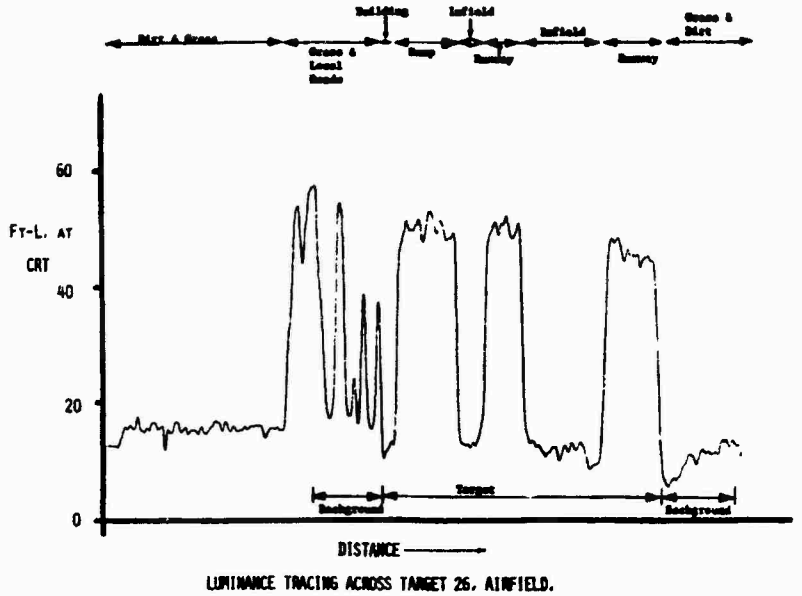



Figure 51. Schematic Representation of Microdensitometric Scan.



LUMINANCE TRACING ACROSS TARGET 26, AIRFIELD.

Figure 52. Microdensitometric Trace of Target 26, Airfield.

Using these traces, a total of 14 photometric measurements was made, several of which were made in two different ways. In conformity with the traditional definition of the background, integration and averaging were performed over that part of the tracing which did not include the target, i.e., A to B and C to D in Figure 51. As an alternate, less exact, but easier technique, the entire width, A to D was defined as the background. This definition, of course, might offer real advantages in an automated scanning system.

Similarly, the area encompassing the distance from 25% of the target's width to the left of the target to 25% of the target's width to the right of the target was defined as the 25% background, either conventionally from points E to B and C to F, or nonconventionally, from point E to point F.

Using either the conventional or nonconventional definition of the background, the following measurements were made from the tracings.

1. INT TOT BKD. Integrated total background is the integrated transmission of the microdensitometer output, frame edge to frame edge for the nonconventional case, and from A to B and from C to D (Figure 51) for the conventional case.
2. INT 25% BKD. The integrated transmission from point E to point F defined this measure for the nonconventional background, while the integrated transmission from point E to point B plus that from point C to point F defined it for the conventional background.
3. CROSS $\bar{X}$ PEAKS. This measure is the number of times the tracing crossed the mean of all maxima and minima, with the mean computed over the entire tracing.
4. CROSS $\bar{X}$ 25%. This measure is the number of times the tracing, between points E and B, and between points C and F, crossed the mean of all maxima and minima between those same pairs of points.
5. $\sigma$ PEAKS. This is the standard deviation of all maxima and minima for the entire tracing.
6. # SIGN CHANGES. The number of sign changes is the number of slope reversals, positive to negative or vice versa, over the entire tracing. Equivalently, it is the number of local maxima plus local minima.
7. INT TGT LUM. Integrated target luminance is the integrated transmission from point B to point C.
8. INT DET LUM. Integrated detail luminance is the integrated transmission from one edge to the other of the detail which was considered most critical to the recognition of the target. This detail was previously specified from an examination of a photograph of the target.
9. MAX TGT LUM. Maximum target luminance is the maximum transmission occurring between points B and C.
10. MAX DET LUM. Maximum detail luminance is the maximum transmission occurring within the bounds of the critical target detail.

11. $\bar{X}$ PEAKS. The mean transmission of all maxima and minima of the entire tracing is termed the mean of the peaks.

12. $\bar{X}$ PEAKS, BKD. This measure is the mean transmission of all maxima and minima between points A and B, and between points C and D.

13. $\bar{X}$ PEAKS 25%. This measure is the mean of all maxima and minima between points E and B, and between points C and F.

14. $\bar{X}$ PEAKS TGT. The mean of all maxima and minima between points B and C was calculated.

PREDICTOR VARIABLES

From these 14 photometric measures, 32 predictors were formed. For simplicity, all predictors with the prefix "A" were based upon the nonconventional definition of the background, while the predictors with the prefixes "B", "C", and "D" used conventional background measurement. Table 13 defines the combination of the above measurements which comprise each predictor variable.

The last three variables in table 13, F1 through F3, are not based upon the microdensitometric scans, but rather upon measured physical sizes of the target detail, respectively (ref. 18). Therefore, a total of 35 predictor variables was used.

CRITERION VARIABLES

Four performance measures were inserted into the prediction program. These were (1) $P_{c-0}$, the proportion of targets correctly recognized at the zero noise level, averaged across all subjects; (2) $P_{c-5}$, the proportion of targets correctly recognized at all five noise levels, averaged across all subjects; (3) $\bar{R}_0$, the mean slant range of recognition at the zero noise level, averaged across all subjects; and (4) $\bar{R}_5$, the mean recognition slant range for all five noise levels, averaged across all subjects.

TABLE 13. COMPOSITION OF PREDICTOR VARIABLES.

| Variable Number | Name | Photometric Measurement Formula | Other Measurement |
|---|---|---|---|
| *A1 | Integrated Target Contrast | (7-1)/1 | |
| *A2 | Integrated Target Contrast, 25% | (7-2)/2 | |
| *A3 | Integrated Target Modulation | (7-1)/(7+1) | |
| *A4 | Integrated Target Modulation, 25% | (7-2)/(7+2) | |
| *A5 | Integrated Detail Contrast | (8-1)/1 | |
| *A6 | Integrated Detail Contrast, 25% | (8-2)/2 | |
| *A7 | Integrated Detail Modulation | (8-1)/(8+1) | |
| *A8 | Integrated Detail Modulation, 25% | (8-2)/(8+2) | |
| B1 | Integrated Target Contrast | (7-1)/1 | |
| B2 | Integrated Target Contrast, 25% | (7-2)/2 | |
| B3 | Integrated Target Modulation | (7-1)/(7+1) | |
| B4 | Integrated Target Modulation, 25% | (7-2)/(7+2) | |
| B5 | Integrated Detail Contrast | (8-1)/1 | |
| B6 | Integrated Detail Contrast, 25% | (8-2)/2 | |
| B7 | Integrated Detail Modulation | (8-1)/(8+1) | |
| B8 | Integrated Detail Modulation, 25% | (8-2)/(8+2) | |
| C1 | Maximum Target Contrast | (9-12)/12 | |
| C2 | Maximum Target Contrast, 25% | (9-13)/13 | |
| C3 | Maximum Target Modulation | (9-12)/(9+12) | |
| C4 | Maximum Target Modulation, 25% | (9-13)/(9+13) | |
| C5 | Maximum Detail Contrast | (10-12)/12 | |
| C6 | Maximum Detail Contrast, 25% | (10-13)/13 | |
| C7 | Maximum Detail Modulation | (10-12)/(10+12) | |
| C8 | Maximum Detail Modulation, 25% | (10-13)/(10+13) | |
| D1 | Mean Target Contrast | (14-12)/12 | |
| D2 | Mean Target Contrast, 25% | (14-13)/13 | |
| D3 | Mean Target Modulation | (14-12)/(14+12) | |
| D4 | Mean Target Modulation, 25% | (14-13)/(14+13) | |
| E1 | Mean Luminance Crossings | 3 | |
| E2 | Mean Luminance Crossings, 25% | 4 | |
| E3 | Standard Deviation, Peaks | 5 | |
| E4 | Number Luminance Reversals | 6 | |
| F1 | Target Size | | Target Length x Target Width |
| F2 | Target Detail Size | | Detail Length x Detail Width |
| F3 | Target Aspect Ratio | | Target Length/ Target Width |

* Denotes use of nonconventional background measurements, as described above.

## RESULTS

Each of the 35 predictor variable measurements was attempted on each of the 25 targets used in the experiment described in detail in section IV of this report. However, at the range chosen for making the microdensitometric scan, four of the targets (Numbers 22, 23, 31, 51) could not be identified on the individual film frame because, at the magnification used with the microdensitometer, the contrast of the target against its background was subliminal. Thus, these four targets were eliminated from the statistical analyses, leaving a total of 21 targets for which the following analyses were made.

A linear stepwise multiple regression program (ref. 24) was used to determine the weighting and importance of each of the predictor variables in predicting, on a target-by-target basis, each of the four criterion variables. Thus, four linear stepwise multiple regression analyses were made. No added (or trans-generated) variables were used, although the Biomedical Program (ref. 24) permits the inclusion of additional predictor variables which are transformations of the initial predictor variables.

In summary, this program determines the intercorrelations among all the variables, predictor and criterion, and then determines which predictor variable best singularly predicts the criterion. It then determines the additional predictor variable which, when added to the first as a multiple regression predictor, most increases the multiple linear correlation, and by what amount. It repeats this step, adding (or deleting) predictor variables in successive steps, to improve the multiple linear correlation until either it is not possible to increase the correlation, until a correlation of unity is obtained, or until the increased prediction falls short of an arbitrarily set criterion of improvement. For our purposes, the criterion was set deliberately liberal so that all variables would be included if they contributed to an increase in the correlation coefficient.

Table 14 shows the summary of this analysis for the prediction of the criterion variable $P_{c-0}$. The analysis summaries for the other three criterion variables are given in tables 15 through 17.

TABLE 14. SUMMARY OF MULTIPLE STEPWISE LINEAR REGRESSION TO PREDICT $P_{c-0}$.

| Step Number | Variable Entered | Multiple R |
|---|---|---|
| 1 | A4 | .5234 |
| 2 | E3 | .6313 |
| 3 | B8 | .7210 |
| 4 | A2 | .7857 |
| 5 | C1 | .8994 |
| 6 | A5 | .9209 |
| 7 | F3 | .9300 |
| 8 | C5 | .9503 |
| 9 | A1 | .9588 |
| 10 | C3 | .9682 |
| 11 | D3 | .9741 |
| 12 | E4 | .9844 |
| 13 | B3 | .9898 |
| 14 | D1 | .9955 |
| 15 | B6 | .9978 |
| 16 | E1 | .9989 |
| 17 | E2 | .9996 |
| 18 | D4 | .9996 |
| 19 | F1 | .9997 |

As indicated in table 14, $P_{c-0}$ can be totally predicted (multiple R = 0.9997) by using a total of 19 of the 35 predictor variables. The first predictor is the fourth variable listed in table 13, variable A4, Integrated Target Modulation, 25%. This variable alone predicts 27.4% of the variability in the criterion variable, as shown by the $R_{SQ}$ value in table 14. Similarly, the addition of variable 31 (E3, Standard Deviation, Peaks) accounts for an additional 12.46% of the variability. Perhaps of most importance is not the last multiple correlation in this table of 0.9997, but rather the fact that 90% of the variability is predicted after only 8 steps, or by the inclusion of only 8 of the 35 predictor variables.

In a similar manner, it can be seen from table 15 that 19 steps are needed to produce a multiple correlation of 1.0000 using $P_{c-5}$ as the criterion variable, and that 90% of the variability is predicted with only 10 steps, using 10 predictor variables.

TABLE 15. SUMMARY OF MULTIPLE STEPWISE LINEAR REGRESSION TO PREDICT $P_{c-5}$.

| Step Number | Variable Entered | Multiple R |
|---|---|---|
| 1 | B1 | .6243 |
| 2 | C5 | .7503 |
| 3 | E2 | .7974 |
| 4 | B8 | .8119 |
| 5 | B5 | .8473 |
| 6 | A3 | .9028 |
| 7 | F1 | .9177 |
| 8 | D3 | .9295 |
| 9 | D1 | .9425 |
| 10 | A8 | .9511 |
| 11 | F3 | .9613 |
| 12 | A6 | .9700 |
| 13 | E4 | .9790 |
| 14 | C1 | .9892 |
| 15 | E1 | .9949 |
| 16 | C3 | .9975 |
| 17 | E3 | .9993 |
| 18 | C4 | .9996 |
| 19 | C2 | 1.0000 |

From table 16, it is seen that 19 steps are needed to produce a multiple correlation of 1.0000, but that only six steps are necessary to predict 90% of the variance of $\overline{R}_0$. Of particular interest is the fact that the first variable (number 28, or D4, Mean Target Modulation, 25%) predicts 48.74% of the variance of $\overline{R}_0$ alone.

TABLE 16. SUMMARY OF MULTIPLE STEPWISE LINEAR REGRESSION TO PREDICT $\bar{R}_0$.

| Step Number | Variable Entered | Multiple R |
|---|---|---|
| 1 | D4 | .6981 |
| 2 | B1 | .8161 |
| 3 | C1 | .8561 |
| 4 | A8 | .8924 |
| 5 | D2 | .9230 |
| 6 | A6 | .9546 |
| 7 | F1 | .9630 |
| 8 | E1 | .9750 |
| 9 | C3 | .9787 |
| 10 | F2 | .9803 |
| 11 | B5 | .9860 |
| 12 | A3 | .9903 |
| 13 | D3 | .9933 |
| 14 | C5 | .9963 |
| 15 | C8 | .9975 |
| 16 | C2 | .9979 |
| 17 | B6 | .9985 |
| 18 | C4 | .9997 |
| 19 | B8 | 1.0000 |

Finally, table 17 illustrates the fact that a multiple correlation of 1.0000 is reached at the 19th step, and that 90% of the variance of $\bar{R}_5$ is predicted by 9 variables.

In summary, then, each of the criterion variables can be p rfectly predicted by 19 or fewer steps in the linear stepwise multiple regression approach, and a maximum of 10 steps or variables is needed to predict 90% of the variance in each of the criterion var bles. The specific variables needed to predict each of the criterion variables varies, of course, but with some commonality, as shown in table 18, which indicates the predictor variables that are needed to

predict 90% or more of the variance for each of the criterion variables, and also the photometric and geometric components of each of these predictor variables.

TABLE 17. SUMMARY OF MULTIPLE STEPWISE LINEAR REGRESSION TO PREDICT $\overline{R}_5$.

| Step Number | Variable Entered | Multiple R |
|---|---|---|
| 1 | B1 | .6452 |
| 2 | A8 | .7338 |
| 3 | D1 | .8185 |
| 4 | A6 | .8865 |
| 5 | E4 | .9034 |
| 6 | C4 | .9169 |
| 7 | C3 | .9248 |
| 8 | B4 | .9373 |
| 9 | D3 | .9550 |
| 10 | A3 | .9637 |
| 11 | B5 | .9727 |
| 12 | E2 | .9775 |
| 13 | B6 | .9866 |
| 14 | C1 | .9898 |
| 15 | E3 | .9941 |
| 16 | E1 | .9990 |
| 17 | F2 | .9994 |
| 18 | C8 | .9999 |
| 19 | C7 | 1.0000 |

The investigator using the linear stepwise multiple regression approach must decide *a priori* what criterion he will employ for inclusion of the next step. One criterion often applied is that the increase in R be significant at, say, $p < .05$. Another popular criterion is that the predicted variance, $R^2$, be increased by at least 5% by each included step. Still other investigators have arbitrarily chosen other, more liberal, criteria. Because this research is of an exploratory nature, and because applied statisticians have not, among themselves, come to any agreement regarding an appropriate criterion, we chose to

let the computer run to a maximum value of R. Thus, the reader can set his own arbitrary criterion, should he wish, and reinterpret the data accordingly.

TABLE 18. PHOTOMETRIC AND GEOMETRIC MEASUREMENTS WHICH, COMBINED, PREDICT 90% OF THE VARIANCE OF CRITERION VARIABLES.

| Regression Predictor Variable | $P_{c-0}$ | $P_{c-5}$ | $\bar{R}_0$ | $\bar{R}_5$ |
|---|---|---|---|---|
| A1 | | | | |
| A2 | 2,7 | | | |
| A3 | | 1,7 | | |
| A4 | 2,7 | | | |
| A5 | 1,8 | | | |
| A6 | | | 2,8 | 2,8 |
| A7 | | | | |
| A8 | | 2,8 | 2,8 | 2,8 |
| B1 | | 2,8 | 2,8 | 2,8 |
| B2 | | | | |
| B3 | | | | |
| B4 | | | | 2,7 |
| B5 | | 1,8 | | |
| B6 | | | | |
| B7 | | | | |
| B8 | 2,8 | 2,8 | | |
| C1 | 9,12 | | 9,12 | |
| C2 | | | | |
| C3 | | | | 9,12 |
| C4 | | | | 9,13 |
| C5 | 10,12 | 10,12 | | |
| C6 | | | | |
| C7 | | | | |
| C8 | | | | |
| D1 | | 12,14 | | 12,14 |
| D2 | | | 13,14 | |
| D3 | | 12,14 | | 12,14 |
| D4 | | | 13,14 | |
| E1 | | | | |
| E2 | | 4 | | |
| E3 | 5 | | | |
| E4 | | | | 6 |
| F1 | | TGT. LxW | | |
| F2 | | | | |
| F3 | TGT L./TGT. W | | | |

In an attempt to determine which of the geometric and photometric measures are most important, the frequency of occurrence of each in table 18 was counted and is summarized in table 19. As shown there, INT 25% BKD, the integrated transmission of the background 25% either side of the target is one of the most important variables, another being INT DET LUM, the integrated transmission of the target detail. Less frequently occurring variables are $\overline{X}$ PEAKS, BKD and $\overline{X}$ PEAKS, TGT, which relate to the mean transmission of the background and the target, respectively.

TABLE 19. FREQUENCY OF USAGE OF INDIVIDUAL PHOTOMETRIC AND GEOMETRIC VARIABLES IN TABLE 18.

| | Variable | Frequency of Usage |
|---|---|---|
| 1 | INT TOT BKD | 3 |
| 2 | INT 25% BKD | 13 |
| 4 | CROSS $\overline{X}$ 25% | 1 |
| 5 | $\sigma$ PEAKS | 1 |
| 6 | # SIGN CHANGES | 1 |
| 7 | INT TGT LUM | 4 |
| 8 | INT DET LUM | 12 |
| 9 | MAX TGT LUM | 4 |
| 10 | MAX DET LUM | 2 |
| 12 | $\overline{X}$ PEAKS, BKD | 9 |
| 13 | $\overline{X}$ PEAKS, 25% | 3 |
| 14 | $\overline{X}$ PEAKS, TGT | 6 |
| TGT | Length | 2 |
| TGT | Width | 2 |

SIMPLIFIED PREDICTION

As more variables are used in a multiple regression prediction equation, there obviously becomes more opportunity to capitalize on chance covariation among the variables and upon sample uniqueness. One way to reduce this tendency is to apply a correction for shrinkage, where the linear stepwise multiple regression coefficient is reduced by an amount expected to eliminate the bias from the

unique sample of data points used, and to offer an estimate of the coefficient which would be obtained if another random sample from the same population were used. The correction for shrinkage is applied by the following formula to obtain a corrected correlation coefficient, $R_s^2$:

$$R_s^2 = 1 - (1-R^2)\left[\frac{n-1}{n-k-1}\right] \qquad (14)$$

where $R_s^2$ = the shrunken multiple correlation squared,

$R^2$ = the multiple correlation squared as obtained from the existing sample,

n = the number of items (e.g., targets) in the sample, and

k = the number of predictors in the regression equation.

In this particular application, all regression coefficients were equal to unity after 19 steps, so that the correction for shrinkage, as determined by equation (14), remains at unity, i.e., the bracketed term equals zero.

Another way to look at a simplified approach to the multiple regression prediction is to reduce the number of predictors to those which seem the most potent. A quick estimate of the value of this screening technique was made by choosing only those 8 predictor variables in table 18 which were used in two or more of the criterion prediction equations. Thus, only variables A6, A8, B1, B8, C1, C5, D1, and D3 were considered. Using only these 8 predictors, the linear stepwise multiple regression program was again run, with the same targets and the same 4 criterion variables. The results are summarized in table 20, which shows that multiple linear correlation coefficients ranging from .724 to .940 are obtained, which predict from 52% to 88% of the criterion variance. Thus, using only 8 photometric measures (the geometric variables are excluded from this list), a large portion of the criterion variance is predicted for each of the four performance measures.

At this point, it seems inappropriate to explore the multiple regression relationships any further. If time permits, the same predictor and criterion variables will be used in subsequent experiments, and similar analyses will be made with

the intent of arriving at some valid prediction equation which is consistent across more targets and other image quality conditions. At present, however, it might be noted that the magnitude of prediction is similar to that reported by Zaitzeff (ref. 21), who used seven predictors and obtained multiple Rs of 0.89 to 0.91. In addition, because of the arbitrary relative transmission units used in the microdensitometry and the transmission-to-video-luminance conversion factors, which are specific to the transilluminance of the film in the projector gate, no coefficients are provided for the variables used in this prediction. In subsequent research, when the multiple prediction equations are better known, such constants and equations will be given.

TABLE 20. MULTIPLE Rs OBTAINED WITH SIMPLIFIED STEPWISE MULTIPLE REGRESSION ANALYSIS.

| | $P_{c-0}$ | | $P_{c-5}$ | | $\overline{R}_0$ | | $\overline{R}_5$ | |
|---|---|---|---|---|---|---|---|---|
| Step | Variable | R | Variable | R | Variable | R | Variable | R |
| 1 | B1+ | .473 | B1+ | .624 | B1+ | .652 | B1+ | .645 |
| 2 | C5+ | .572 | C5+ | .750 | C1+ | .842 | A8+ | .734 |
| 3 | B8+ | .606 | A8+ | .774 | A8+ | .883 | D1+ | .819 |
| 4 | A3+ | .655 | A6+ | .805 | B8+ | .904 | A6+ | .887 |
| 5 | A6+ | .683 | C1+ | .819 | D1+ | .915 | C5+ | .890 |
| 6 | D3+ | .707 | B8+ | .822 | D3+ | .936 | C1+ | .892 |
| 7 | B1- | .707 | D3+ | .822 | A6+ | .940 | | |
| 8 | C5- | .706 | D1+ | .823 | C5+ | .940 | | |
| 9 | D1+ | .717 | | | | | | |
| 10 | B1+ | .721 | | | | | | |
| 11 | C5+ | .723 | | | | | | |
| 12 | C1+ | .724 | | | | | | |

+ denotes entering variable

- denotes removing variable

## SECTION VII

## GENERAL DISCUSSION AND CONCLUSION

The preceding sections of this report have presented the results of several experimental studies, with the generally consistent conclusions that both $MTFA_{SQ}$ and the averaged $SNR_D$ are excellent predictors of the differences in performance across line-scan system configurations. It was also pointed out that the $N_e$ concept has little or no application to these data because it is insensitive to differences in system noise levels, especially when noise is independent of the video bandpass of the imaging system. In the following paragraphs, an attempt will be made to compare, analytically, these different image quality metrics, to summarize the results thus far in this program, and to describe a conceptual model which combines system and scene parameters to predict observer performance.

### COMPARISON OF MTFA AND $SNR_D$

Reference 2 has recently compared MTFA and $SNR_D$. That analysis, modified to meet the terminology of this application, shows that there are certain similarities between MTFA and $SNR_D$ from which one might conclude that they become equally valid predictors of operator performance.

Equation (5) of this report stated that the eye's theoretical threshold detection requirement for a sinusoidally varying periodic intensity pattern of frequency N on a static photograph is:

$$M_t(N) = 0.034 \left( \frac{dD}{d(\log_{10} E)} \right)^{-1} \left( 0.033 + \sigma(D)^2 N^2 S^2 \right)^{1/2} \qquad (5)$$

in which N = any spatial frequency, in lines per millimeter

0.034 = an empirically derived constant

D = mean film density

E = exposure

0.033 = an empirically derived constant

$\sigma(D)$ = rms granularity for a 24μ scanning aperture

S = signal-to-noise ratio necessary for threshold viewing, assumed to be about 4.5.

The constant 0.033 represents the limitation of the eye to very low spatial frequency inputs, to which the eye is limited in its spatial integration capability, or its "DC" response. For purposes of this derivation it can be ignored. Then, assuming gamma, or $dD/d(\log_{10}E)$, to be unity, this equation becomes

$$M_t(N) = 0.034\ \sigma(D)\ N\ S \tag{15}$$

The term $\sigma(D)$ is the rms granularity or noise in the photograph, which in video terms is simply rms noise; the term N can be expressed in lines per picture height rather than in lines per millimeter; and S can be considered analogous to the threshold signal-to-noise ratio for a 50% probability of detection. By changing the constant 0.034 to $\beta$ to reflect the change in units, above, equation (15) becomes

$$M_t(N_{TV}) = \beta\ i_n\ N_{TV}\ SNR_{DT} \tag{16}$$

where $i_n$ is the noise photocurrent,
$SNR_{DT}$ is the threshold $SNR_D$, and
$N_{TV}$ is the number of lines per picture height

Rearranging, this equation becomes

$$SNR_{DT} = \frac{1}{\beta} \cdot \frac{M_t(N_{TV})}{N_{TV}} \cdot \frac{1}{i_n} \quad . \tag{16a}$$

Letting $M_t(N_{TV})$ equal $\Delta i_{ST}$, the threshold video signal,

$$SNR_{DT} = \frac{1}{\beta} \cdot \frac{1}{N_{TV}} \cdot \frac{\Delta i_{ST}}{i_n} \quad . \tag{17}$$

However, $\Delta i_{ST}/i_n$ is the threshold signal-to-noise ratio in the video, as used by Rosell in his analyses (e.g., ref. 3), so that equation (17) becomes

$$SNR_{DT} = \frac{(1/\beta)\ SNR_{\text{video, threshold}}}{N_{TV}} \quad . \tag{18}$$

In equation (7), it was stated that the linear MTFA was defined by

$$MTFA = \int_0^{N_1} \left( T_N - \frac{M_t(N)}{M_o} \right) dN \quad . \tag{7}$$

In the video system, the object's inherent contrast at the photosurface can be represented by $Ci_s$, the image contrast times the highlight photocurrent. Therefore, using equation (16) in equation (7),

$$MTFA = \int_0^{N_1} \left( T_N - \frac{\beta i_n N_{TV} SNR_{DT}}{Ci_s} \right) dN \quad . \tag{19}$$

Now $Ci_s/i_n$ is the broad-area video signal-to-noise ratio that the sensor can produce at unity contrast input conditions. Letting $Ci_s/i_n$ equation $SNR_{V,0}$,

$$MTFA = \int_0^{N_1} \left( T_N - \frac{\beta N_{TV} SNR_{DT}}{SNR_{V,0}} \right) dN \tag{20}$$

$$= \int_0^{N_1} \left( \frac{T_N \cdot SNR_{V,0} - \beta SNR_{DT} N_{TV}}{SNR_{V,0}} \right) dN \quad . \tag{20a}$$

By substituting, as per equation (18),

$$MTFA = \int_0^{N_1} \left( \frac{T_N \cdot SNR_{V,0} - SNR_{video,\ threshold}}{SNR_{V,0}} \right) dN \quad . \tag{21}$$

However, $T_N$ is the sine-wave response at $N_{TV}$ lines per picture height, and is equal to the signal-to-noise ratio at that line number, $SNR_{V,N}$, divided by the signal-to-noise ratio at N=0, or the DC response of the system. That is,

$$SNR_{V,N} = T_N \cdot SNR_{V,0} \quad . \tag{22}$$

Therefore, combining equations (21) and (22),

$$MTFA = \int_0^{N_1} \left( \frac{SNR_{V,N} - SNR_{video,\ threshold}}{SNR_{V,0}} \right) dN \quad . \qquad (23)$$

Equation (23) is the integral of the difference between the signal-to-noise ratio at $N_{TV}$ lines per picture height and the eye's signal-to-noise requirement at threshold, normalized to the video signal-to-noise ratio at N=0. The integration is performed from N equals zero to the value of N at which $SNR_{V,N}$ equals $SNR_{video,\ threshold}$. The resulting integral, illustrated in figure 53(b), is equal to the MTFA (figure 53(a)) under certain conditions.

Specifically, this derivation is valid <u>if and only if</u> the system responses are measured using a sinusoidal input, the system gamma is unity, and the visual threshold requirement is determined for a sinusoidal target. It also assumes ideal viewing conditions, display magnification, and viewing time. Note that the experiments described in previous sections of this report did not use sinusoidal inputs, which is the reason for employing $MTFA_{SQ}$ as the notation rather than MTFA.

However, by the time the square-wave input is passed through the system, with its limiting aperture response, the stimulus to the eye is, for all practical purposes, a sine-wave intensity pattern except at very low spatial frequencies. At higher spatial frequencies, the visual thresholds to sine and square waves are very similar (ref. 26). Further, one would not expect substantial differences in correlation between the MTFA and the $MTFA_{SQ}$ values with observer performance if the conditions producing variation in observer performance were reasonably large. Thus, for all practical purposes, the use of the $MTFA_{SQ}$ metric in this research and the integral of the $SNR_D$ metric, across all applicable spatial frequencies, should produce the same prediction accuracy when large system configuration differences are evaluated.

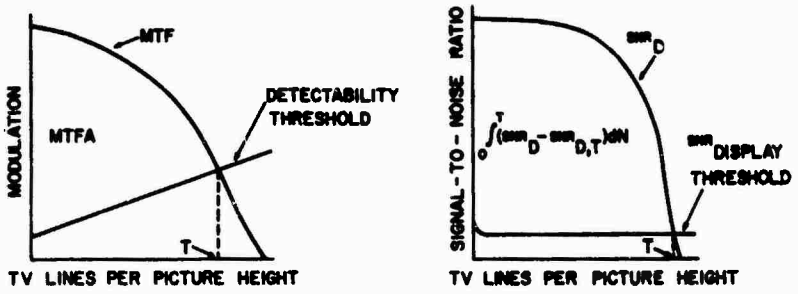



Figure 53. Comparison of $SNR_D$ and MTFA.

$$MTFA = \int_0^T (SNR_D - SNR_{D,T})\, dN \text{ under conditions stated in text.}$$

This is not the same as saying, however, that $SNR_D$ equals MTFA, but merely that an integral, based upon $SNR_D$, is equal to MTFA. The application of $SNR_D$ by Rosell and Willson (ref. 3) includes a discrete value of $SNR_D$ for the specific target spatial frequency and system under consideration, and does not include the integral of $SNR_D$ minus a threshold level over a range of spatial frequencies. This difference is extremely important when one considers that the usual airborne reconnaissance application is one in which the target of interest comes into the field of view at some distance (at a high spatial frequency) and gradually approaches the observer until it passes out of the bottom of the displayed image (at a lower spatial frequency). If one wished to know the independent likelihood of finding the target at a particular range (with no information accumulated by the observer until the target reached that range),

then the specific spatial frequency, N, at which $SNR_D$ is calculated would be a perfectly meaningful predictor of total system performance. If, however, one wan to use an $SNR_D$-type measure and is interested in the overall cumulative likelihood of the observer recognizing the target at any time while the target is in the field of view (equivalently, at any spatial frequency), it seems most appropriate to integrate the value of $SNR_D$ (or the value of $SNR_D$ minus some constant threshold) over all spatial frequencies of use, which is simply a value linearly proportional to MTFA. Statistically speaking, the integral over all spatial frequencies (MTFA) is unbiased as to scene content and scene dynamics, whereas the discrete spatial frequency $SNR_D$ is necessarily specific to a given target magnification, although it can be calculated for all magnifications.

To date, research by Rosell and his associates has shown good prediction of detection and recognition performance using static, non-time-limited scenes in which the size of the target does not change during a single trial. To the best of our knowledge, the only application of the $SNR_D$ concept, albeit a modified version, to dynamic imagery is that used in section IV of this report, and that is merely an averaged value over many scenes. Thus, the application of $SNR_D$ to the detection or recognition of a target, the magnification and viewing aspect of which are changing, requires additional analysis, perhaps akin to the integration approach used in the $MTFA_{SQ}$ model.

In this context, of course, the $SNR_D$ metric can be used to predict the range at which a given target's size is large enough to cause a threshold $SNR_D$ value. Thus, the $SNR_D$ approach can be used to predic. the specific range at which the target can be detected (or recognized, or discriminated), but any spread about that specific range is then calculated from an assumed gaussian distribution of probabilities (see ref. 3). We know of no experimental validation of this concept for dynamic imagery to date, but would predict that the same type of result would be obtained using such an $SNR_D$ approach with some empirically determined probability density function, as has been obtained with the $MTFA_{SQ}$ metric.

Related to this argument is the fact that any imaging system is typically used for a myriad of purposes beyond those intended by the designer. Thus, using a concept of evaluation ($MTFA_{SQ}$) which contains, with equal weights, all spatial frequencies which the system can image is statistically unbiased and most generalizable. If one is interested in system performance at only a single spatial frequency, however, either the value of $SNR_D$ or the difference between $R_{sq}(N)$ and the detectability threshold at that spatial frequency (section III) can be used. In reality, however, real scenes viewed during the performance of real tasks contain a wide variety of spatial frequencies. A concept for handling this problem is presented later in this section.

PERIODIC VS. NONPERIODIC TARGET THRESHOLDS

In a recent paper, Schade (ref. 27) advocated the use of a signal-to-noise measure in which the threshold value needed for detection is based upon the mean of that for a periodic (e.g., three-bar) target and that for a single, nonperiodic (e.g., one-bar) target. The assumption in his paper, undoubtedly valid, is that the world is composed also of nonperiodic targets, so that a threshold requirement based upon only periodic targets is artificial. Schade's assumption is certainly appropriate; however, its application can be questioned in terms of the necessary level of refinement in order to achieve a statistically maximum prediction. That is, when one is predicting the inherently variable performance of a nonlinear system, such as the human observer, only a certain maximum degree of prediction is obtainable, and persons doing research on human behavior have, for many years, in even the most rigorously controlled, laboratory environment, been content, often deliriously happy, with correlations on the order of .90 or better. To attempt to define a metric of image quality which is significantly more predictive than that shown for the $MTFA_{SQ}$ may be a fruitless, even naive search. Thus, while Schade's argument is undoubtedly theoretically ideal, it may be a case of analytical overkill, especially since no human performance data are presented to support it.

Further, the equal weighting of periodic and nonperiodic targets in specifying threshold signal-to-noise requirements is also arbitrary - to the best of our knowledge, the real world is composed more of nonperiodic elements, assuming it were necessary to make the distinction.

It should also be noted that the qualitative evaluation of MTFA, as indicated by Schade (pp. 569-570, ref. 33) is based upon a misinterpretation of the MTFA concept. Schade assumes, in his analysis, that the threshold function varies in slope with changes in exposure, although no source is referenced in his paper. However, as section III of this report has shown, the slope of the threshold curves, when required modulation is plotted as a function of spatial frequency, is invariant with noise levels, and no mention need be made of the exposure value, since the displayed modulation takes into account the input modulation and the gamma of the system. That is, one can specify the required display modulation at threshold independently of exposure, although the knowledge of exposure and gamma is needed to determine the modulation ultimately produced by the system for a given target input.

## INDIVIDUAL TARGET PREDICTION

There is no doubt, from the data reported in these experiments and others, that the prediction of an observer's ability to detect or recognize a specific target on a specific display under specific operating conditions is very difficult. Certainly, merely knowing the MTF or the $R_{sq}(N)$ of the system, the target's size and shape, the noise level of the system, and the mean luminance contrast of the entire target with its background is not nearly enough. As discussed in section VI, other factors are also related, such as the luminance of the target detail, the highlight luminance, the variability of luminance in the background, and the photometric characteristics of the local (25%) background. None of the image quality measures proposed thus far can handle this multitude of parameters. However, an intuitively reasonable approach, using the combination of spatial frequency analysis of the scene and the system MTF, is suggested, as follows.

Assume first that the display contains two sources of information, one originating from the scene and another originating from the imaging system. Also, assume that the information can be divided into vertical and horizontal components, or components perpendicular to the raster and components parallel to the raster when the raster is oriented in the commercially common horizontal direction. The reason for this distinction will become clear later. Looking first at the horizontal (parallel to the raster) dimension, a certain amount of system

dynamic noise is likely to be present, with a particular power spectral density distribution, such as that shown in figure 54(a). All noise frequencies are not of equal importance (section III), however, and some weighting function must be applied to emphasize the lower frequencies of dynamic noise (ref. 25). Letting this weighting function be of the form shown in figure 54(b), the effective displayed noise is that shown in figure 54(c).

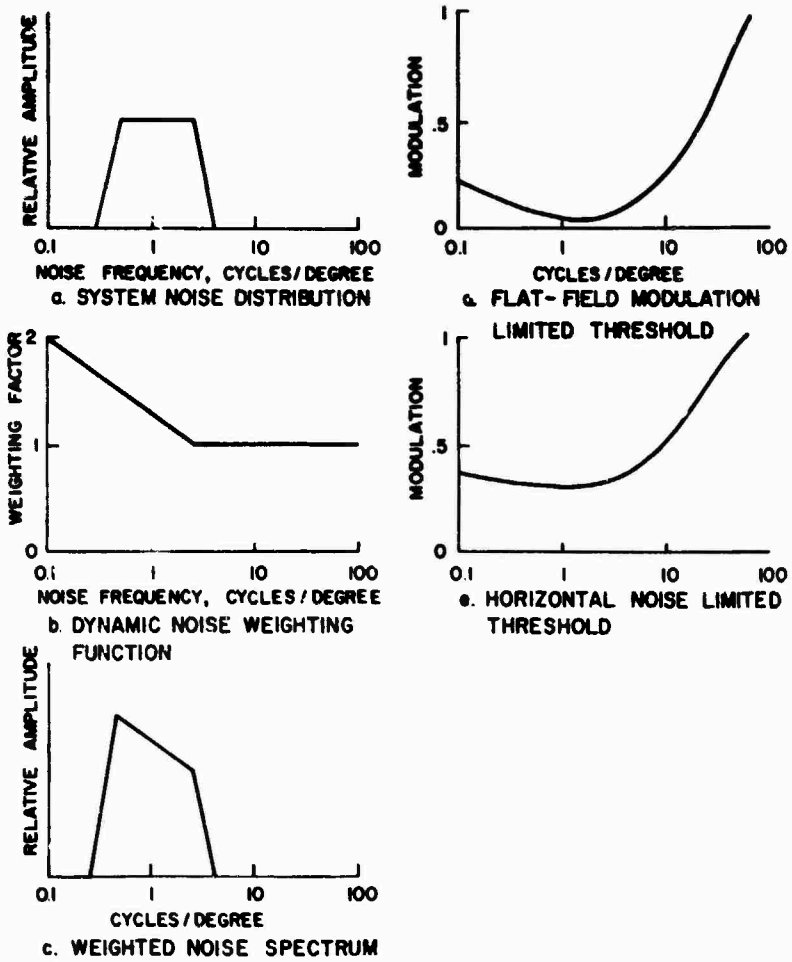



Figure 54. Horizontal Components Analysis.

Were there no dynamic noise at the display, the flat field, modulation-limited threshold would be of the form given in figure 54(d) (ref. 13). The addition of noise causes the threshold to increase as some function of the noise spectrum, with the noise effect ranging from one octave below the noise passband to at least one octave above the noise passband (ref. 26), although the extent and degree of such threshold modification for the line-raster display is not well defined at present. Presently on-going research in this laboratory will define this effect for representative noise passbands; for the purposes of the present discussion, however, assume that the dynamic noise-limited threshold is of the form given in figure 54(e).

Turning now to the vertical (perpendicular to raster) dimension, the dominant system noise source is the raster frequency, which can have a serious interference effect upon detecting objects with spatial frequency components close to the raster frequency (ref. 27). Figure 55(a) shows a representative spectrum for the raster. Because the raster is a static (noise) pattern, no weighting function is needed, as is the case for the horizontal noise spectrum. The flat-field threshold of the eye is virtually the same in the vertical dimension as in the horizontal, but is necessarily affected by the raster noise to produce a noise-limited threshold for this dimension, as shown in figure 55(b) which combines the flat-field, modulation limited threshold with the raster interference effect. The dynamic noise (figure 54(a)) will also have some effect in the vertical dimension, although it will be somewhat uneven due to the discrete raster sampling. Letting figure 55(c) represent a modified dynamic noise threshold in this dimension, then figure 55(d) can be used to combine the dynamic and static noise thresholds in the vertical dimension. The means by which these thresholds are combined, as well as the precise definition of these thresholds, is under study and must be defined empirically. For purposes of this conceptual model, at least, figures 54(e) and 55(d) can be taken to represent the horizontal and vertical noise-limited thresholds, respectively.

Figure 56(a) contains an assumed scene power spectral density distribution (ref. 28), referenced to the angular field of view of the imaging system. This scene spectrum is passed through the horizontal MTF of the system (figure 56(b)), resulting in the display horizontal spectrum of the scene, figure 56(d), in

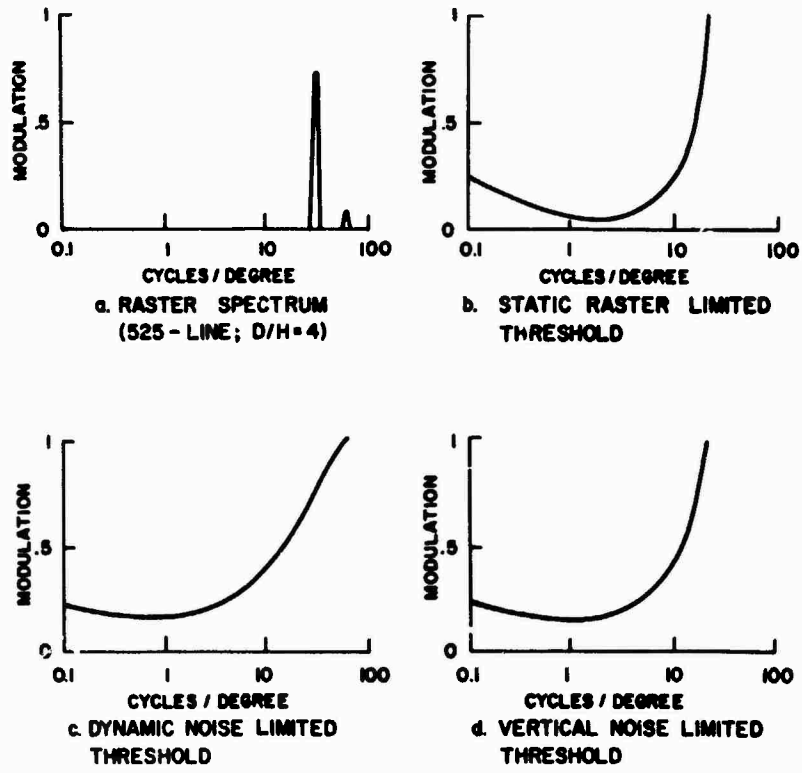



Figure 55. Vertical Components Analysis.

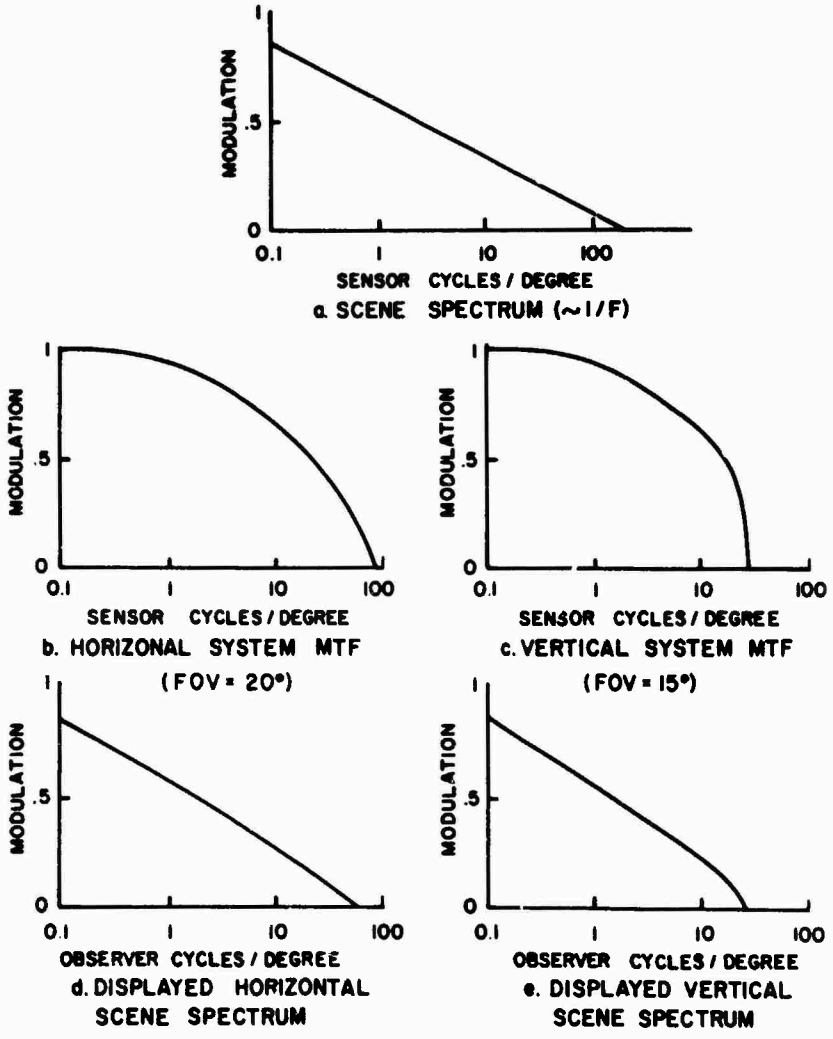



Figure 56. Displayed Vertical and Horizontal Scene Spectrum.

observer angular units. Similarly, the scene spectrum is passed through the vertical MTF of the system (figure 56(c)) to yield a displayed vertical scene spectrum figure 56(e). An isotropic system will yield equal vertical and horizontal MTFs, but that is not to say that it is desirable to have an isotropic system. Previous research has shown that a vertical raster is preferable to a horizontal raster, even if the system is isotropic (ref. 29); present research in this laboratory is investigating the performance of observers using anisotropic MTFs.

The scene power density spectra, taken separately for the vertical and horizontal dimensions, are cross-plotted with the vertical and horizontal noise-limited thresholds in figure 57. The cross-hatched area represents the degree to which the scene spectrum exceeds the threshold detectability spectrum of the observer. Note that in figure 56 all plots are in terms of cycles per degree, referenced to the observer. As display size and viewing distance change, these curves will obviously shift.

The extent to which the target is detectable (or recognizable or identifiable) in the displayed scene is a function of the excess displayed scene spectrum over the displayed noise threshold, or the cross-hatched areas in figures 57(a) and 57(b). Note, however, that there is no reason to expect that this area is directly proportional to the likelihood of detectability or to any other measure of observer performance. Rather, it is the case that the target cannot be detected if it does not exceed this noise-limited threshold; whether it will be detected is a function of many other task- and observer-related variables.

For example, the target and background spectra are combined in figure 56(a). The observer does not distinguish between target and nontarget objects until he detects the target. Thus, a target may be undetected simply because its spectrum is not distinguishable from its local background (section VI), or because the observer simply does not look in the right location. Present experiments are comparing eye-movement data for various types of display conditions, both static and dynamic, to relate eye-movement distributions and search strategies to scene content. In addition to scene content and associated spectra, the observer may also view the display with a predetermined "set" or cognitive map which causes

him to scan the display in an uneven manner and thereby not look at the target. Such effects will always remain a source of unpredictable variance in any visual search data, although it is possible to train persons to scan a display more uniformly (ref. 30).

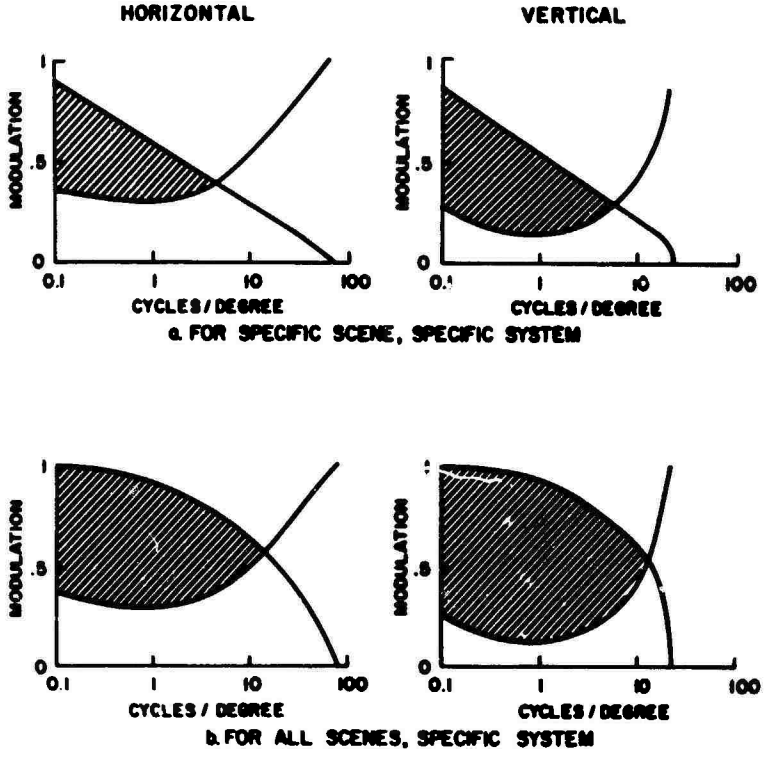



Figure 57. Excess of Displayed Scene Spectrum Over Visual Threshold.

As the scene movement rate changes (e.g., as with decreased altitude, increased look-down angles, and increased groundspeed), the system MTF may change slightly, leading to a blurred image. Also, the angular rates may become so high as to decrease the observer's dynamic visual acuity (ref. 31), although such rates are typically excessive for airborne displays. The major effect of increased display rate of motion is to limit search time, rather than to degrade the image. A reduction in search time necessarily reduces the number of fixations the observer can make in search for the target while it is within the field of view, and may force him to alter his eye fixation distribution to a less-than-optimal pattern. Additional data are needed to quantify this relationship.

No attempt has been made here to combine the vertical and horizontal measures of scene content in excess of thresholds (figure 57), although data are presently on hand to do so. Considerable more analysis is required at this time, and will be presented in a future report under this contract. Candidate means for combinations of the vertical and horizontal values include (1) simple summation and (2) converting both into a volumetric measure. The former approach would assume independence between the vertical and horizontal components, which is not likely; the latter approach becomes somewhat more complex, and the resulting quantity is not as heuristically desirable.

Several advantages accrue to this conceptual approach. First, it admits variable spectra for the target/background combination, noise sources, and system elements, including the raster interference pattern. Although raster effects have been demonstrated in terms of preferred viewing distances, no data have been published to date to show that the raster interferes with information extraction. Such an experiment has recently been completed in this laboratory, with the data suggesting that raster interference effects are not nearly as great as anticipated (ref. 27). The results of this experiment will be contained in a separate report.

Second, the model weights noise frequencies by their effect upon the visual system, and distinguishes between static and dynamic noise types. It has been suggested that all system noise is not white. If this be true, then the unequal effect of various noise bands (section III) makes it all the more important to evaluate the noise effect upon the detection threshold in terms of its power density spectrum.

Third, the inclusion of the system MTF retains the analytical convenience of predicting system performance early in the design process, given any scene spectrum and theoretical system noise spectrum. The approach further permits evaluation of the tradeoff between horizontal and vertical system response, once the best means is determined by which the cross-hatched areas of figures 57(a) and 57(b) are combined.

Fourth, the inclusion of the scene spectrum permits a given system, real or hypothetical, to be evaluated for performance against any scene content. Input spectra are obviously affected by certain system and mission variables, e.g., terrain, ground cover, field of view, altitude, speed, spectral sensitivity, optics speed, scene irradiance, etc.

The authors are aware of certain limitations of this conceptual model as well. For example, it assumes that the human observer's decision-making system is a linear one which operates largely upon the spatial frequency content of the display. This notion is quite simplistic, although it may serve as an adequate approximation for the task of form recognition as it has in other tasks (ref. 32). The model also largely disregards instructions to the observer, eye movement scan efficiency, cognitive maps of the functional relationships in the search scene, etc. While ongoing research in this laboratory is related to some of these unanswered questions, it must be realized that there will always remain a measurable variance in human form recognition or target acquisition which can never be predicted, due simply to inherent variability among observers. As demonstrated in section III, such unpredicted variance cannot be reduced below about 12% for the simple tri-bar detection task. Thus, there is no reason to expect better prediction for a more complex visual task. By the same logic, it is critical that researchers in this area use appropriate inferential statistics to estimate the remaining unpredicted variance in order that their analytical models can be evaluated against some appropriate estimate of utility. Mere plotting of mean performance data is typically inadequate and misleading.

## REFERENCES


1. Schade, O. H. Image gradation, graininess and sharpness in television and motion-picture systems. Part III: The grain structure of television images. J. SMPTE, 1953, 61, 97-164.

2. Biberman, L. M., Rosell, F. A., Schnitzler, A. D., Schade, O. H., and Snyder, H. L. Low-light-level devices: A designer's manual. IDA Report R-169, August 1971.

3. Rosell, F. A. and Willson, R. H. Performance synthesis (Electro-Optical Sensors). USAF Report AFAL-TR-71-137, May 1971.

4. Middleton, W. E. K. Vision through the atmosphere. Univ. of Toronto Press, 1952.

5. Corbett, D. G., Diamantides, N. D., and Kause, R. H. Measurements and models for relating the physical characteristics of images to target detection. USAF Report AMRL-TR-64-117, December 1964.

6. Jennings, L. B., Meeker, F. B., Praver, G. A., and Cook, R. N. Ground resolution study final report. USAF Report RADC-TR-63-224, 29 November 1963.

7. Brainard, R. W., Sadacca, R., Lopez, L. J., and Ornstein, G. N. Development and evaluation of a catalog technique for measuring image quality. USAPRO Report 1150, August 1966.

8. Charman, W. N. and Olin, A. Image quality criteria for aerial camera systems. Photog. Sci. Engin., 1965, 9, 385-397.

9. Borough, H. C., Fallis, R. F., Warnock, T. H., and Britt, J. H. Quantitative determination of image quality. Boeing Report D2-114058-1, May 1967.

10. Scott, F. In Image evaluation for reconnaissance, Proceedings of a symposium. Itek Report 9048-6, April 1963.

11. Carman, P. D. and Charman, W. N. Detection, recognition and resolution in photographic systems. J. Opt. Soc. Amer., 1964, 54, 1121-1130.

12. Campbell, F. W. and Green, D. G. Optical and retinal factors affecting visual resolution. J. Physiol., 1965, 181, 576-593.

13. DePalma, J. J. and Lowry, E. M. Sine-wave response of the visual system. II. Sine-wave and square-wave contrast sensitivity. J. Opt. Soc. Amer., 1962, 52, 328-335.

14. Lowry, E. M. and DePalma, J. J. Sine-wave response of the visual system. I. The Mach phenomenon. J. Opt. Soc. Amer., 1961, 51, 740-746.

15. Bryngdahl, O. Characteristics of the visual system: psychophysical measurements of the response to spatial sine-wave stimuli in the photopic region. J. Opt. Soc. Amer., 1966, 56, 811-821.

16. Biberman, L. M. Display size, brightness level and observer response time. Paper presented to the Society for Information Display, Boston, October 1966.

17. Klingberg, C. L., Elworth, C. L., and Filleau, C. R. Image quality and detection performance of military interpreters. Final report of USAFOSR Contract F 44620-69-C-0128, Boeing Company, April 1970.

18. Humes, J. M. and Bauerschmidt, D. K. Low light level TV viewfinder simulation program. Phase B Report. USAF Report AFAL-TR-68-271, November 1968.

19. Scott, F. The relationship between three-bar target resolving power and modulation transfer functions. Perkin-Elmer paper (unpublished).

20. Coltman, J. W. and Anderson, A. E. Noise limitations to resolving power in electronic imaging. Proceedings of the IRE , 1960, 858-865.

21. Zaitzeff, L. P. Target background scaling and its impact on the prediction of aircrew target acquisition performance. Boeing Technical Report D180-14156-1, December 1971.

22. Rhodes, F. Predicting the difficulty of locating targets from judgments of image characteristics. USAF Report AMRL-TDR-64-19, March 1964.

23. Nygaard, J. E., Slocum, G. K., Thomas, J. O., Skeen, J. R., and Woodhull, J. G. The measurement of stimulus complexity in high-resolution sensor imagery. USAF Report AMRL-TDR-64-29.

24. Dixon, W. J. (Ed.) BMD Biomedical Computer Programs, Berkeley: U. of Calif. Press, 1970.

25. Schober, H. A. W. and Hilz, R. Contrast sensitivity of the human eye for square-wave gratings. J. Opt. Soc. Amer., 1965, 55, 1086-1091.

26. Stromeyer, C. F., III, and Julesz, B. Spatial-frequency masking in vision: Critical bands and spread of masking. J. Opt. Soc. Amer., 1972, 62, 1221-1232.

27. Schade, O. H. Image reporduction by a line raster process. In. L. M. Biberman (Ed.) Perception of displayed information. Plenum Press, 1973.

28. Scott, F. and Frauenhofer, D. The modulation transfer function and methods of measurement. In L. M. Biberman and S. Nudelman (Eds.) Photoelectronic Imaging Devices, Vol. 1. Plenum Press, 1971, pp. 291-306.

29. Rusis, G. Laboratory studies on air-to-ground target recognition: VIII. The effect of TV image enhancement in the observer-initiated freeze mode. Autonetics Report T6-276/3111, 14 January 1966.

30. Overby, L. T. The application of adaptive training techniques to visual search. Unpublished Ph.D. dissertation, Virginia Polytechnic Institute and State University, November, 1973.

31. Snyder, H. L. and Greening, C. P. The effect of direction and velocity of relative motion upon dynamic visual acuity. Autonetics Report C5-477/3111, 31 January 1965.

32. Cornsweet, T. N. Visual perception, Academic Press, 1970.

33. Schade, O. H. Resolving power functions and integrals of high definition television and photographic cameras - a new concept of image evaluation. RCA Review, 1971, 32, 567-609.

## Appendix A. Wideband Video Mixer

The two-input video mixer (Figure A-1) is constructed on a printed circuit board using discrete components and conventional techniques.

The two input transistors are connected as a differential pair. The 32 MHz video signal is fed to one, and the 20 MHz (or less) white noise to the other. The sum, signal plus noise, appears at the collectors. This mixed signal is fed to two transistors which form the line driver. The video input and output impedance is 75 ohms, and the noise input impedance is 52 ohms. The output voltage is flat to 32 MHz for ranges between 0.4 and 1.6 volts. A 15% pre-emphasis has been added to the output at the high frequency end of the spectrum to compensate for high frequency cable loss. Three voltage levels are provided by an external power supply.

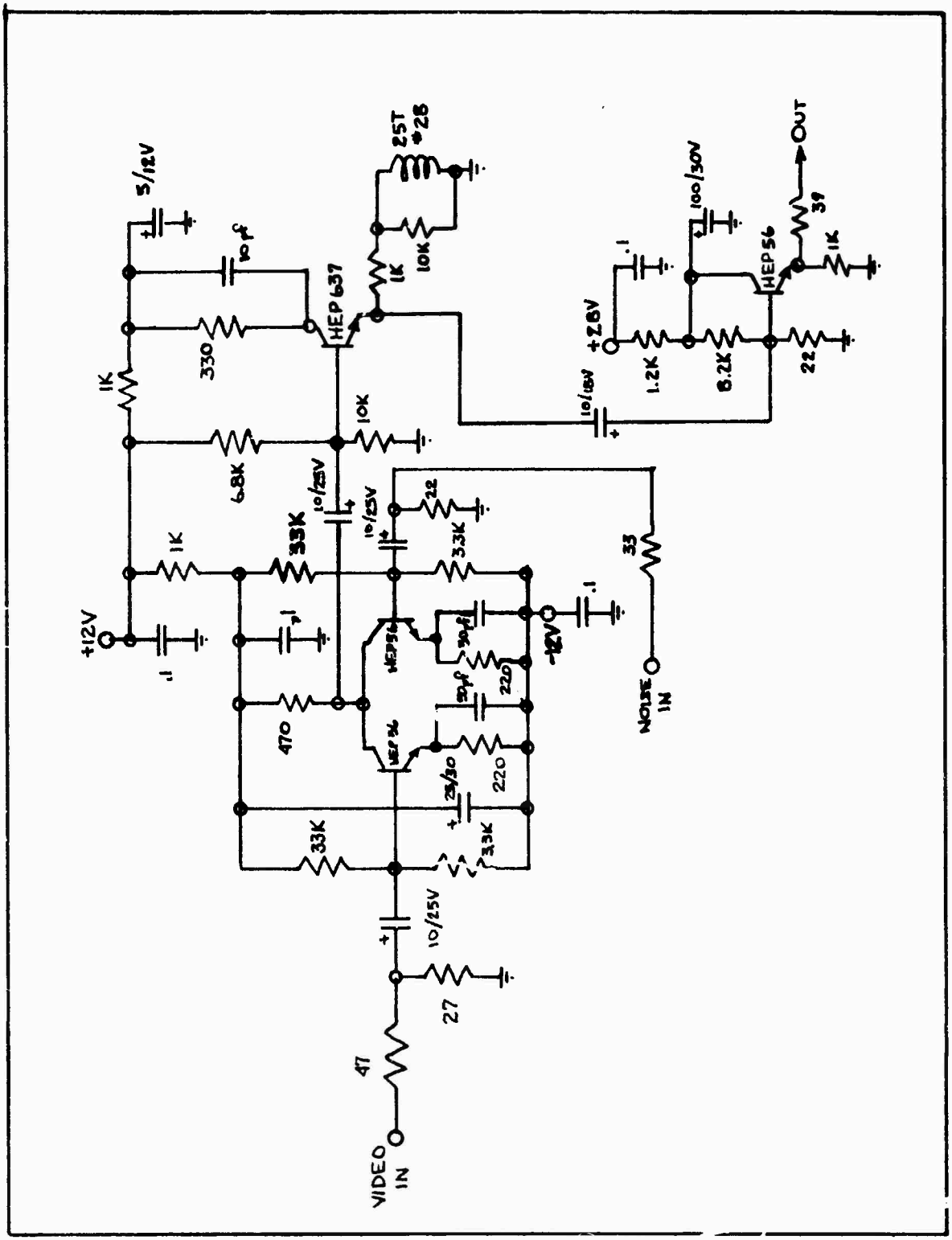



Figure A-1. Video Mixer Schematic Diagram.

## Appendix B. Sync Generator/Frame Counter

The synchronization processing unit (Figure B-1) generates the sync signals necessary to lock the video system and the Strobex to the projector. It also automatically switches the video camera control unit to line sync when the projector is not running. An adjustable sync delay is provided to the Strobex unit in order to fire it at the proper time with respect to the projector shutter and video camera scan.

The sync processor consists of several sets of circuits. All are SN 7400 series TTL with the exception of a 10 V. discrete circuit to provide eight-volt sync pulses to the Strobex.

One-half of an SN 7413 dual Schmitt trigger provides a 60 Hz square wave necessary for the system to be locked to the AC line frequency in the absence of projector sync.

The other half of the SN 7413 detects the presence of camera sync and switches the system sync from line to projector. The projector sync is divided by two and counted with a five decade counter. The projector or line sync is then fed to a monostable multivibrator which produces a delayed sync for the Strobex unit with respect to the camera control unit.

The frame counter (Figure B-2) consists of five decade counting units, each with a latch memory and seven segment light emitting diode display. Each film frame is counted and displayed, and may be frozen (latched) by the subject by means of a hand-held pushbutton. The experimenter provides the unlatch input after recording the particular frame number. Pushbutton bounce conditioning is provided, and the counter has provisions for BCD output to a printer for automatic recording.

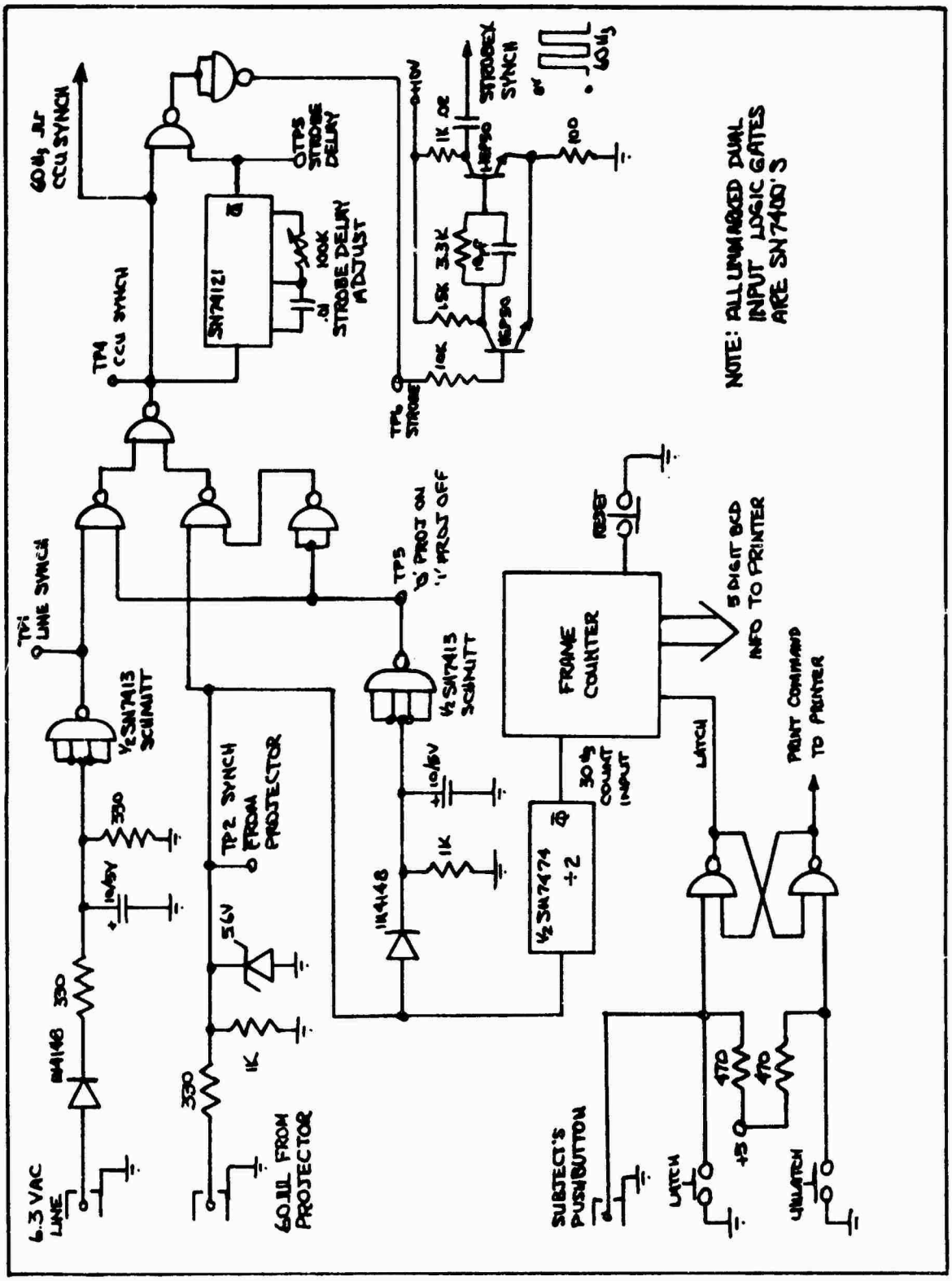



Figure B-1. Synchronization Processing Unit.

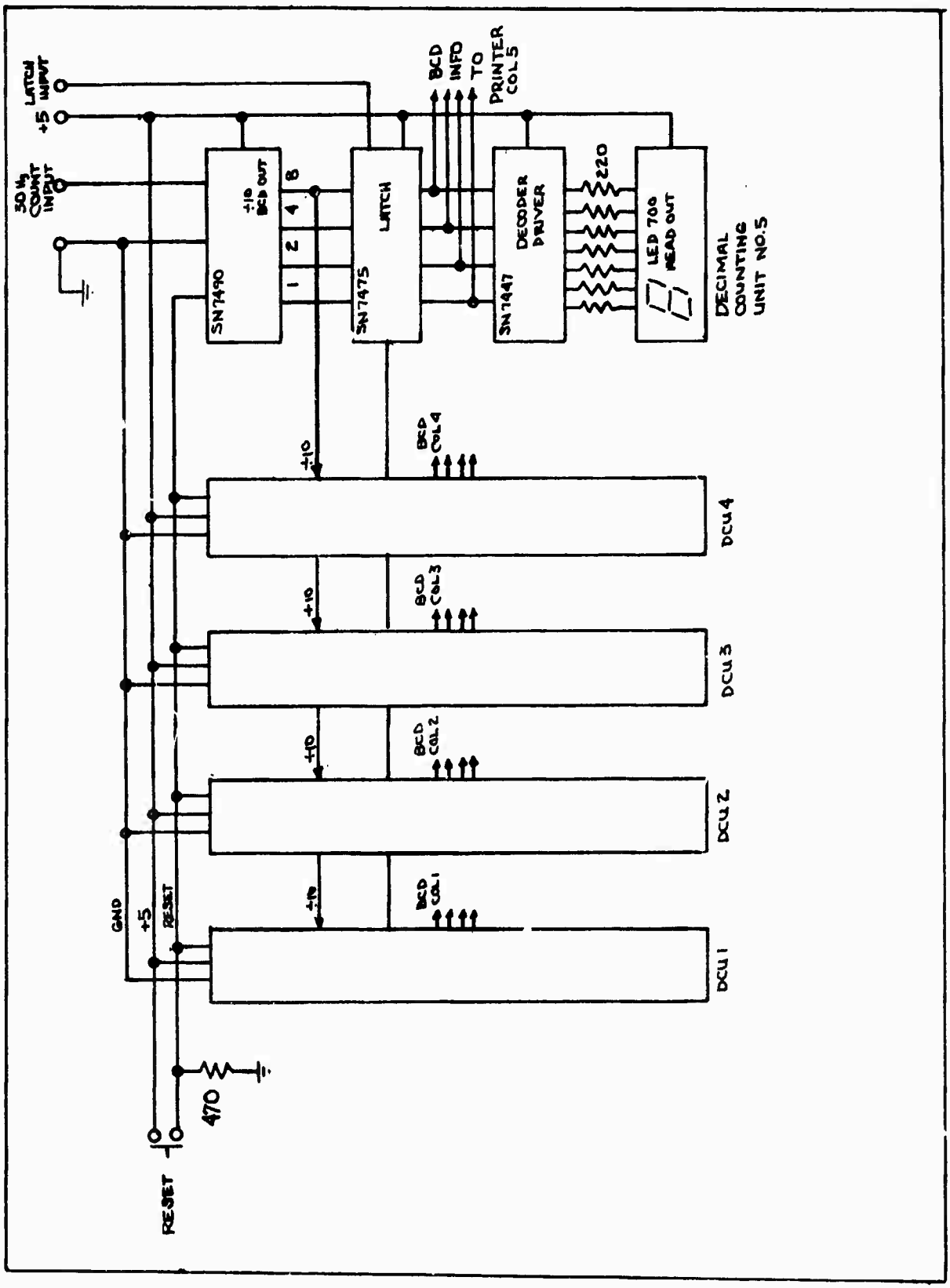



Figure B-2. Frame Counter Schematic Diagram.



*U.S.Government Printing Office: 1975 — 657-018/403